

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:57:12 ; Search time 4329.72 Seconds
(without alignments)
16865.305 Million cell updates/sec

Title: US-08-731-499-3
Perfect score: 1507
Sequence: 1 GCAGTGTCTGGATTGACT.....GGTCTTGCACTCCAGGGTCC 1507

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pac.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1504.8	99.9	1507	6	AR070326 Sequence
2	1504.8	99.9	1507	6	BD085727 Genes fro
3	1387.2	92.1	10365	6	BD085733 Genes fro
4	1387.2	92.1	105023	2	AC116668 Trypanoso
5	1387.2	92.1	121143	9	AF312915 Homo sapi
6	1387.2	92.1	128871	9	AL157838 Human DNA
7	1337.2	88.7	5632	6	AR157086 Sequence
8	1337.2	88.7	5632	9	AF041259 Homo sapi
9	1337.2	88.7	5633	6	CQ726295 Sequence
10	1108.6	73.6	3186	6	BD085734 Genes fro
11	614.2	40.8	184420	2	AC118373 Rattus no
12	614.2	40.8	246332	2	AC099079 Rattus no
13	573.6	38.1	114285	10	AL844576 Mouse DNA
14	573.6	38.1	208979	2	AC023610 Mus muscu
15	217.8	14.5	3512	5	BC072191 Xenopus 1
16	217.4	14.4	2750	5	AF419155 Xenopus 1
17	197.2	13.1	298	6	AX895965 Sequence
18	197.2	13.1	298	6	BD031498 Sequence
19	156.4	10.4	469	6	BD060160 Secreted

20	114.4	7.6	212026	2	AC136666	Rattus no	
21	114.4	7.6	244851	2	AC094607	Rattus no	
C	22	114.4	7.6	268559	2	AC112808	Rattus no
23	105.4	7.0	4934	10	BC053104	Mus muscu	
C	24	105.4	7.0	175110	10	AC142100	Mus muscu
25	104.8	7.0	5982	6	CQ714295	Sequence	
C	26	104.8	7.0	112797	9	AC018413	Homo sapi
C	27	104.8	7.0	178717	2	AC024396	Homo sapi
28	104.8	7.0	211981	9	AC009716	Homo sapien	
29	103.2	6.8	6033	9	D86975	Homo sapien	
30	97	6.4	8511	5	BC081149	xenopus 1	
C	31	79.4	5.3	165094	2	CR547124	Danio rer
32	79.4	5.3	174225	5	BX005071	Zebrafish	
C	33	79.4	5.3	192578	5	AL732455	Zebrafish
C	34	79.4	5.3	243887	2	AC024898	Homo sapi
35	78.6	5.2	4265	10	BC033594	Mus muscu	
36	78.6	5.2	4417	10	AK122270	Mus muscu	
37	78.6	5.2	195922	10	AC140213	Mus muscu	
38	74.6	5.0	256281	2	AC094062	Rattus no	
C	39	74.6	5.0	266956	2	AC133822	Rattus no
C	40	73.2	4.9	171239	2	AC150136	Gallus ga
41	73.2	4.9	190305	2	AC150054	Gallus ga	
42	70	4.6	70	6	CQ876260	Sequence	
43	70	4.6	4935	6	CQ718121	Sequence	
44	70	4.6	4935	9	AB002388	Homo sapi	
45	70	4.6	163120	9	AC011454	Homo sapi	

ALIGNMENTS

RESULT 1
AR070326
LOCUS AR070326 1507 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 3 from patent US 5892010.
ACCESSION AR070326
VERSION AR070326.1 GI:7221214
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1507)
AUTHORS Gray,J., Collins,C., Hwang,S.-i., Godfrey,T., Kowbel,D. and
Rommens,J.
TITLE Genes from the 20Q13 amplicon and their uses
JOURNAL Patent: US 5892010-A 3 06-APR-1999;
FEATURES Location/Qualifiers
source
1..1507
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN	Query Match	99.9%;	Score 1504.8;	DB 6;	Length 1507;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 1507;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	GCAGTGTCTGGATTGACTTCTTCTCAATTTGAACACTCATTCAATGGACACAAAGAG	60		
Db	1	GCAGTGTCTGGATTGACTTCTTCTCAATTTGAACACTCATTCAATGGACACAAAGAG	60		
Qy	61	CACATAATGCTTTGCTGCTGATTCATATTTGAATTCGAGGCATTTGGAAACCTCTGATGCTTGG	120		
Db	61	CACATAATGCTTTGCTGCTGATTCATATTTGAATTCGAGGCATTTGGAAACCTCTGATGCTTGG	120		
Qy	121	TTTGTGGAAGAACACGATGACACCACTCACTGAGCTTCTTAAAGTTCGAGAGATTAGAG	180		
Db	121	TTTGTGGAAGAACACGATGACACCACTCACTGAGCTTCTTAAAGTTCGAGAGATTAGAG	180		
Qy	181	GACTATACACTTCTTTTGAACCTTTTATAATAAATTTTGTCTCTGCTTGTGTTGAAACCCAGG	240		
Db	181	GACTATACACTTCTTTTGAACCTTTTATAATAAATTTTGTCTCTGCTTGTGTTGAAACCCAGG	240		
Qy	241	ACTGTTAGAGGGTGTAGTGACAGGTCTTACAGTGGGCTTAATCAAACCTCCAGAAATTTGCC	300		

```
Db 241 ACTGTTAGAGGTGAGTCACAGTCTTACAGTGGCTTAATCCAACTCCAGAAATTGCC 300
Qy 301 AACGGAACCTTTGAGATTATATGCAATCGAAATGACAGGAACATGCCAACTCAATCCCT 360
Db 301 AACGGAACCTTTGAGATTATATGCAATCGAAATGACAGGAACATGCCAACTCAATCCCT 360
Qy 361 CTTAATGTACATGGATGCCAAGAGTGATGGCAGCTCTCTTGCCAGTCCGATCGAGATG 420
Db 361 CTTAATGTACATGGATGCCAAGAGTGATGGCAGCTCTCTTGCCAGTCCGATCGAGATG 420
Qy 421 GAGATGCCCTTGTCAATGAAGAGGCCNCTGTTGTCAATTCGAGCTACACAAAGAAAAA 480
Db 421 GAGATGCCCTTGTCAATGAAGAGGCCNCTGTTGTCAATTCGAGCTACACAAAGAAAAA 480
Qy 481 ATGTCATTCGAAATCGAGGGGAATATGCCCCCTTGGATTGCAATGTTCTGAGCCAGACCTTC 540
Db 481 ATGTCATTCGAAATCGAGGGGAATATGCCCCCTTGGATTGCAATGTTCTGAGCCAGACCTTC 540
Qy 541 ACACATTCAGAGACCTTAAATAAATGCTTAAATGCAACACCGGCTACCTCTGTGAA 600
Db 541 ACACATTCAGAGACCTTAAATAAATGCTTAAATGCAACACCGGCTACCTCTGTGAA 600
Qy 601 CCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGGA 660
Db 601 CCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGGA 660
Qy 661 ACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGATGTGGG 720
Db 661 ACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGATGTGGG 720
Qy 721 CAGACATTTAGAGTCGCTTTTGTATGCTGAGATCCACATGAGAACACACAAAGATCTTTC 780
Db 721 CAGACATTTAGAGTCGCTTTTGTATGCTGAGATCCACATGAGAACACACAAAGATCTTTC 780
Qy 781 ACTTACGGGTGTAAATGTGCGGAAAGAAATGCAAGGAGCTTGTGTTTCTTAAATACAC 840
Db 781 ACTTACGGGTGTAAATGTGCGGAAAGAAATGCAAGGAGCTTGTGTTTCTTAAATACAC 840
Qy 841 ATGCGGACRCATATGCGAAATCGGGGGCCAGAGCAAACTGCAGCAAGGCTTGAGAGT 900
Db 841 ATGCGGACRCATATGCGAAATCGGGGGCCAGAGCAAACTGCAGCAAGGCTTGAGAGT 900
Qy 901 AGTCAGCAAGCATCAACGAGTCTGTCAGGTGTCAGCGGCGGAGAGCATCTCTCTCT 960
Db 901 AGTCAGCAAGCATCAACGAGTCTGTCAGGTGTCAGCGGCGGAGAGCATCTCTCTCT 960
Qy 961 TGCATAATCTGATGGTTTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAATGAGCAC 1020
Db 961 TGCATAATCTGATGGTTTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAATGAGCAC 1020
Qy 1021 CGCAAGGTGCACACAAAAAATCTGTTTGGTACAGAGCGGCGGAGAGCTCTTCCA 1080
Db 1021 CGCAAGGTGCACACAAAAAATCTGTTTGGTACAGAGCGGCGGAGAGCTCTTCCA 1080
Qy 1081 CAAGGAGAAATGCGCTCTCGAGGGAGGACTTCTGCAAGTGTCTCACTTGAGACCAAAA 1140
Db 1081 CAAGGAGAAATGCGCTCTCGAGGGAGGACTTCTGCAAGTGTCTCACTTGAGACCAAAA 1140
Qy 1141 TCTCAACCTGAAACGGGGAAGAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTACC 1200
Db 1141 TCTCAACCTGAAACGGGGAAGAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTACC 1200
Qy 1201 ACCTTCAGGCTTGGCAKCTGCTACCAAGAAWAGTTGCCATTTGCCAAGAGTGAAG 1260
Db 1201 ACCTTCAGGCTTGGCAKCTGCTACCAAGAAWAGTTGCCATTTGCCAAGAGTGAAG 1260
Qy 1261 GAAATGGGGAAGAGGAGGACCCACACACGAGTTCGAGTTCGAGAGGAGCTTGA 1320
Db 1261 GAAATGGGGAAGAGGAGGACCCACACACGAGTTCGAGTTCGAGAGGAGCTTGA 1320
Qy 1321 GAAACAAATAAGAACCATTTGTGAGGCTCTCGCAAGAGAAAGAGAGTGCACACCTCC 1380
Db 1321 GAAACAAATAAGAACCATTTGTGAGGCTCTCGCAAGAGAAAGAGAGTGCACACCTCC 1380
```

```
1321 GAAACAAATAAGAACCATTTGTGAGGCTCTCGCAAGAGAAAGAGAGTGCACACCTCC 1380
Qy 1381 CACGGCAGAGCCCTCGTGGACCGGATCCCAAGTTACCACTAGCAAGAGAGAGCC 1440
Db 1381 CACGGCAGAGCCCTCGTGGACCGGATCCCAAGTTACCACTAGCAAGAGAGAGCC 1440
Qy 1441 ACTCACTGCTCCGAGTGGGCAAGCTTTGAGAACCTTACCACTAGCTGCTTTGCACTCC 1500
Db 1441 ACTCACTGCTCCGAGTGGGCAAGCTTTGAGAACCTTACCACTAGCTGCTTTGCACTCC 1500
Qy 1501 AGGGTCC 1507
Db 1501 AGGGTCC 1507

RESULT 2
BD085727 1507 bp DNA linear PAT 27-AUG-2002
LOCUS Genes from the 20q13 amplicon and their uses.
DEFINITION BD085727
ACCESSION BD085727
VERSION BD085727.1 GI:22631337
KEYWORDS JP 2001524802-A/3.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1507)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and
Rommens,J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 3 04-DEC-2001;
COMMENT THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS Artificial Sequence
PN JP 2001524802-A/3
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11, C12Q1/68, A61K48/00
CC Description of Artificial Sequence:cc49
CC N is A, G, T, or U
FH Key Location/Qualifiers
FT modified base (447).
FEATURES
source
1..1507
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 99.9%; Score 1504.8; DB 6; Length 1507;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCAGTGTCTGGATTGACTTCTTGCTCAATTGAAACACTCATTCATTAAGGACAAAGAG 60
Db 1 GCAGTGTCTGGATTGACTTCTTGCTCAATTGAAACACTCATTCATTAAGGACAAAGAG 60
Qy 61 CACTAATCTTTGTGCTGATTCAATTTGAATTCGAGGCACTGGGAACCTGTATGCTTG 120
Db 61 CACTAATCTTTGTGCTGATTCAATTTGAATTCGAGGCACTGGGAACCTGTATGCTTG 120
Qy 121 TTTGTGGAAGAACACAGTGACACCATCTCTGCTCTCTTATAATAATTTGCTCTGTTTGGAA 180
Db 121 TTTGTGGAAGAACACAGTGACACCATCTCTGCTCTCTTATAATAATTTGCTCTGTTTGGAA 180
Qy 181 GACTATACACTTCTTTTGAACCTTTTATAATAATTTGCTCTGTTTGGAA 240
Db 181 GACTATACACTTCTTTTGAACCTTTTATAATAATTTGCTCTGTTTGGAA 240
Qy 241 ACTGTTAGAGGTGAGTGACAGGCTTTTACAGTGGCTTTAATCCAACTCAGAAATTTGCC 300
```



```
QY 2 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTTCAATGGAGACAAAGAGC 61
Db 6708 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTTCAATGGAGACAAAGAGA 6767
QY 62 ACTAATGCTTTTGCTGATTTCATATTTGAATCGAGGCATTGGGAACCTGTATGCCCTTGT 121
Db 6768 ACTAATGCTTTTGCTGATTTCATATTTGAATCGAGGCATTGGGAACCTGTATGCCCTTGT 6827
QY 122 TTGTGGAAGAAACCACTGACACCATCACTGAGCTTCTTAAAGTTTCGAAGAAGTTAGAGG 181
Db 6828 TTGTGGAAGAAACCACTGACACCATCACTGAGCTTCTTAAAGTTTCGAAGAAGTTAGAGG 6887
QY 182 ACTATACACTTTCTTTTGAACCTTTTATAATAATAATTTGCTCTGG-TTTTTGAACCCAGG 240
Db 6888 ACTATACACTTTCTTTTGAACCTTTTATAATAATAATTTGCTCTGGTTTGTGAACCCAGG 6947
QY 241 ACTGTTTAGA-GGGTAGTGACAGCTCTTAC-AGTGGCCTTTAATCCAACTCCAGAAATTGC 298
Db 6948 GCTGTTTAGAGGGTAGTGACAAAGTCTTACAGTGGCCTTTAATCCAACTCCAGAAATTGC 7007
QY 299 CCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAACAATGCCAACTCAATCC 358
Db 7008 CCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAACAATGCCAACTCAATCC 7067
QY 359 CTCCTTAATGTACATGAT-GGCCAAGAGTGATTTGGCAGCTCTCTTGCCAGTCCGATGGAG 417
Db 7068 CTCCTTAATGTACATGATGGCCAGAAAGTGATTTGGCAGCTCTCTTGCCAGTCCGATGGAG 7127
QY 418 ATGGA-GATGCTCTGTCAATGAAGGGCCNCTGTTGTCAATTCGAGCTACACAAGAA 476
Db 7128 ATGGAAGATGCTCTGTCAATGAAGGGCCNCTGTTGTCAATTCGAGCTACACAAGAA 7187
QY 477 AAAAAATGTCATCCGAATCGAGGGGAATATGCCCTTTGGATTGATGTTCTGAGCCAGAC 536
Db 7188 AAAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTTGGATTGATGTTCTGAGCCAGAC 7246
QY 537 CTTTACACATTCAGAAAGACTTAATAAATGCTTTAATGCAACACCGGCCCTACCTCTG 596
Db 7247 CTTTACACATTCAGAAAGACTTAATAAATGCTTTAATGCAACACCGGCCCTACCTCTG 7306
QY 597 TGAACAGCAGTCTTCCGGTTGAACGAGGATATCTCAGTCCGCTTGATTAAGTCAACT 656
Db 7307 TGAACAGCAGTCTTCCGGTTGAACGAGGATATCTCAGTCCGCTTGATTAAGTCAACT 7366
QY 657 GCGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATG 716
Db 7367 GCGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATG 7426
QY 717 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGATGCCATGATGCCAGAACACAAAGATTC 776
Db 7427 TGGGCAGACATTTAGAGTCGCTTTTGATGTTGATGCCATGATGCCAGAACACAAAGATTC 7486
QY 777 TTTCACTTACGGGTGTAAATGTCGGAAGAGATTCAGAGGACCTTGGTTCTTAAAAA 836
Db 7487 TTTCACTTACGGGTGTAAATGTCGGAAGAGATTCAGAGGACCTTGGTTCTTAAAAA 7546
QY 837 TCACATGCGGACRCAATAATGGCAAAATCGGGGGCAGAACTGCAGCAGGCTTTGA 896
Db 7547 TCACATGCGGACACATAATGGCAAAATCGGGGGCAGAACTGCAGCAGGCTTTGA 7606
QY 897 GAGTAGTCAGCAACCATCAACAGAGTCGTCAGAGGTGCA CGGGCCGAGAGCATCTCTC 956
Db 7607 GAGTAGTCAGCAACCATCAACAGAGTCGTCAGAGGTGCA CGGGCCGAGAGCATCTCTC 7666
QY 957 TCGTTGCAAAATCTGATGTTTGTGGCTTCTTATTTCCAAATAAAGAAGTCTAATTGA 1016
Db 7667 TCGTTGCAAAATCTGATGTTTGTGGCTTCTTATTTCCAAATAAAGAAGTCTAATTGA 7726
QY 1017 GCACCGCAAGGTGCACACCAAAAAAATCTGTTTCGTTACAGCAGCGCGCAGACACTC 1076
Db 7727 GCACCGCAAGGTGCACACCAAAAAAATCTGTTTCGTTACAGCAGCGCGCAGACACTC 7786
QY 1077 TCCACAAGGAGGAATGCCGCTCTCGAGGGAGGACTTCTCTGCAATTTGTTCAACTTGAGACC 1136
```

```
Db 7787 TCCACAAGGAGGAATGCCGCTCTCGAGGGAGGACTTCTCTGCAATTTCAACTTGAGACC 7846
QY 1137 AAAATCTCACTCCCTGAAAACGGGGAAGAGCTGTGCAGATGCATCCCTCAGCTCGATCCGTT 1196
Db 7847 AAAATCTCACTCCCTGAAAACGGGGAAGAGCTGTGCAGATGCATCCCTCAGCTCGATCCGTT 7906
QY 1197 CACCACTTCCAGGCTTGCAKCTGCTGCTACCAAGAAAGTTCATTTGCCAAGAAGT 1256
Db 7907 CACCACTTCCAGGCTTGCGAGCTGGCTACCAAGGAAAGTTGCCATTTGCCAAGAAGT 7966
QY 1257 GAAGNAATTTGGGGCAGAGGAGGACCGACACGACGATTCGAGTTCCGAGAGGAGCT 1316
Db 7967 GAAGNAATCGGGGCAAGAAAGGAGGACCGACACGACGATTCGAGTTCCGAGAGGAGCT 8026
QY 1317 TCGAGAAACAAATTAAGAACCATTTGTGAGGCTCTCGCAAGAGAAAGAGTGCAAAACA 1376
Db 8027 TCGAGNAACAAATTAAGGCGAGTTGTGAGGCTCTCGCAAGAGAAAGAGTGCAAAACA 8086
QY 1377 CTCCACGCGGAGCGCCCTCCGTGGAACGCGATCCCAAGTTACCCAGTAGCAAGAGAA 1436
Db 8087 CTCCACGCGGAGCGCCCTCCGTGGAACGCGATCCCAAGTTACCCAGTAGCAAGAGAA 8146
QY 1437 GCCCACTACTGCTCGAGTGGCGGCAAGCTTTTCAGAACTACCACAGCTGCTCTTGA 1496
Db 8147 GCCCACTACTGCTCGAGTGGCGGCAAGCTTTTCAGAACTACCACAGCTGCTCTTGA 8206
QY 1497 CTCAGGGTCC 1507
Db 8207 CTCAGGGTCC 8217

RESULT 5
AP312915/c
LOCUS AP312915 121143 bp DNA linear PRI 13-JUN-2001
DEFINITION Homo sapiens chromosome 20 clones 97 and 127, complete sequence.
ACCESSION AP312915
VERSION AP312915.1 GI:11094030
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 121143)
AUTHORS Collins,C., Volik,S., Kowbel,D., Ginsinger,D., Ylestra,B.,
Cloutier,T., Hawkins,T., Predki,P., Martin,C., Wernick,M.,
Kuo,W.L., Alberts,A. and Gray,J.W.
Comprehensive genome sequence analysis of a breast cancer amplicon
Genome Res. 11 (6), 1034-1042 (2001)
JOURNAL
MEDLINE 21275464
PUBMED 11381030
REFERENCE 2 (bases 1 to 121143)
AUTHORS Volik,S., Collins,C., Gray,J., Wernick,M., Kowbel,D., Stultz,K. and
Martin,C.
Direct Submission
Submitted (10-OCT-2000) Cancer Genetics, UCSF Cancer Center, 2340
Sutter St., Rm. S151, San Francisco, CA 94706, USA
FEATURES
source
1..121143
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="20"
/maps="20q13.2"
/clone="97; 127; LBNL h117 and h79"
/note="Isolated in Collins lab and sequenced by the Genome
Center at LBL"
complement(1..66093)
/note="overlaps Homo sapiens chromosome 20 clone
RP5-823G15 deposited in GenBank Accession Number
AL354993.19"
misc_feature
194..468
repeat_region
/rpt_family="Alu"
```

repeat_region	/rpt type=dispersed complement(513. .653) /rpt_family="MER5" /rpt_type=dispersed 1053. .1153 /rpt_family="MER20" /rpt_type=dispersed 1254. .1330 /rpt_family="Alu" /rpt_type=dispersed complement(1356. .1676) /rpt_family="Alu" /rpt_type=dispersed 1681. .1866 /rpt_family="Alu" /rpt_type=dispersed 2485. .5421 /note="similar to Homo sapiens chromosome 5 clone CTC-269P1 deposited in GenBank Accession Number AC011341.2" 2594. .2879 /rpt_family="Alu" /rpt_type=dispersed complement(2896. .4437) /note="similar to flow-sorted chromosome 20 HindIII fragment, SC20pF17C6 deposited in GenBank Accession Number Z94882.1" 3106. .3412 /rpt_family="Alu" /rpt_type=dispersed 3889. .4206 /rpt_family="Alu" /rpt_type=dispersed 4449. .4764 /rpt_family="Alu" /rpt_type=dispersed complement(5671. .5987) /rpt_family="Alu" /rpt_type=dispersed 6723. .7004 /rpt_family="Alu" /rpt_type=dispersed complement(7486. .7625) /rpt_family="L1" /rpt_type=dispersed 7716. .7965 /rpt_family="Alu" /rpt_type=dispersed complement(8590. .118606) /note="similar to Mus musculus clone CT7-510C4 deposited in GenBank Accession Number AC023610.3" complement(9852. .10109) /note="similar to STS WI-31010 deposited in GenBank Accession Number G24374.1" <10244. .>10904 /product="hypothetical" /note="similar to Homo sapiens hypothetical protein OV10-2 deposited in GenBank Accession Numbers AF230201.1, AP000365 and in NM 018692.1" /evidence=not experimental 15301. .15588 /rpt_family="Alu" /rpt_type=dispersed 16873. .17073 /note="similar to Homo sapiens STS WI-16697 deposited in GenBank Accession Number G21490.1" 17577. .17857 /rpt_family="Alu" /rpt_type=dispersed complement(19034. .19333) /rpt_family="Alu" /rpt_type=dispersed 20136. .20719	repeat_region repeat_region repeat_region repeat_region repeat_region misc_feature repeat_region misc_feature misc_feature mrna repeat_region misc_feature repeat_region repeat_region repeat_region	/rpt_family="Alu" 20776. .21071 /rpt_family="Alu" /rpt_type=dispersed complement(22344. .22638) /rpt_family="Alu" /rpt_type=dispersed 22726. .22784 /rpt_family="Alu" /rpt_type=dispersed complement(22984. .23268) /rpt_family="Alu" /rpt_type=dispersed complement(23349. .23646) /rpt_family="Alu" /rpt_type=dispersed complement(24008. .24189) /rpt_family="MIR" /rpt_type=dispersed complement(24502. .24795) /rpt_family="Alu" /rpt_type=dispersed 24883. .24941 /rpt_family="Alu" /rpt_type=dispersed complement(25140. .25425) /rpt_family="Alu" /rpt_type=dispersed complement(25507. .25805) /rpt_family="Alu" /rpt_type=dispersed complement(26170. .26352) /rpt_family="MIR" /rpt_type=dispersed 26841. .27129 /rpt_family="Alu" /rpt_type=dispersed 27270. .33290 /rpt_family="L1" /rpt_type=dispersed 28512. .28730 /rpt_family="MER25" /rpt_type=dispersed 34498. .35009 /rpt_family="MLT2B2" /rpt_type=dispersed 35035. .35332 /rpt_family="Alu" /rpt_type=dispersed 35343. .35522 /rpt_family="Alu" /rpt_type=dispersed 35593. .121143 /note="overlaps Homo sapiens clone RP4-724E16 on chromosome 20 deposited in GenBank Accession Number AL157838.24" 36137. .36250 /rpt_family="Alu" /rpt_type=dispersed 36366. .36690 /rpt_family="Alu" /rpt_type=dispersed 36949. .37020 /rpt_family="MER5" /rpt_type=dispersed complement(37692. .37766) /rpt_family="MER5" /rpt_type=dispersed complement(38006. .38276) /rpt_family="Alu" /rpt_type=dispersed 38863. .38984
---------------	---	--	---

[illegible]

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, ENBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/c_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

This sequence is the entire insert of clone RP4-724E16 The true right end of clone RP5-823G15 is at 19684 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP4-724E16 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

FEATURES

source

Location/Qualifiers

1..128871
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="20"
 /map="q13.12-13.32"
 /clone="RP4-724E16"
 /clone_lib="RPCI-4"

repeat_region 4810..5101
 /note="AluSx repeat: matches 1..297 of consensus"
 repeat_region 5110..5413
 /note="AluSx repeat: matches 1..305 of consensus"
 repeat_region 5492..5648
 /note="L2 repeat: matches 2589..2745 of consensus"
 repeat_region 5659..6135
 /note="MERIA repeat: matches 1..488 of consensus"
 repeat_region 6164..6225
 /note="31 copies 2 mer aa 71% conserved"
 repeat_region 6245..6379
 /note="AluSg/x repeat: matches 162..296 of consensus"
 repeat_region 6380..6537
 /note="79 copies 2 mer aa 60% conserved"
 repeat_region 6558..6967
 /note="AluSg repeat: matches 1..309 of consensus"
 repeat_region 7320..7542
 /note="MER58B repeat: matches 117..341 of consensus"
 repeat_region 7545..7838
 /note="AluSx repeat: matches 1..293 of consensus"
 misc_feature 8465..9068
 /note="match: GSS: Em:AQ386102"
 repeat_region 8523..8814
 /note="AluJb repeat: matches 1..294 of consensus"
 repeat_region 9262..9561
 /note="AluSg repeat: matches 1..301 of consensus"
 repeat_region 9562..9599
 /note="19 copies 2 mer tt 81% conserved"
 repeat_region 9606..9733
 /note="MER33 repeat: matches 17..148 of consensus"
 repeat_region 9761..9915
 /note="MER5A repeat: matches 4..183 of consensus"
 repeat_region 9980..10108
 /note="L2 repeat: matches 2622..2747 of consensus"
 repeat_region 10128..10437
 /note="AluSx repeat: matches 2..312 of consensus"
 repeat_region 10654..10724
 /note="MIR repeat: matches 63..137 of consensus"
 repeat_region 10781..11083
 /note="AluSx repeat: matches 1..306 of consensus"
 misc_feature complement(10897..11319)
 /note="match: GSS: Em:AQ499673"
 misc_feature complement(10991..11319)
 /note="match: GSS: Em:AQ283579"
 repeat_region 11462..11759
 /note="AluSx repeat: matches 1..299 of consensus"
 repeat_region 11878..12182
 /note="AluSx repeat: matches 1..300 of consensus"
 repeat_region 12183..12587
 /note="L1PA8 repeat: matches 5757..6163 of consensus"
 repeat_region 13333..13637
 /note="MLTIG repeat: matches 3..338 of consensus"
 repeat_region 14201..14507
 /note="AluY repeat: matches 1..308 of consensus"
 repeat_region 14540..14694
 /note="L2 repeat: matches 2550..2709 of consensus"
 repeat_region 14764..14868
 /note="MIR repeat: matches 103..208 of consensus"
 repeat_region 14891..15152
 /note="AluJo repeat: matches 12..299 of consensus"
 repeat_region 15162..15293
 /note="L2 repeat: matches 1697..1824 of consensus"
 repeat_region 15317..15544
 /note="L1ME3 repeat: matches 5943..6173 of consensus"
 repeat_region 16306..16603
 /note="AluSc repeat: matches 1..306 of consensus"
 repeat_region 16610..16901
 /note="AluSx repeat: matches 1..292 of consensus"
 repeat_region 16989..17035
 /note="MLTIG repeat: matches 420..463 of consensus"
 repeat_region 17036..17338
 /note="AluJo repeat: matches 1..301 of consensus"
 repeat_region 17339..17730
 /note="13 copies 2 mer gt 96% conserved"

repeat_region /note="MUT1C repeat: matches 1. .420 of consensus"
17958. .18269
repeat_region /note="AluX repeat: matches 1. .311 of consensus"
18288. .18315
repeat_region /note="MER91 repeat: matches 32. .59 of consensus"
18383. .18507
repeat_region /note="MIR repeat: matches 23. .148 of consensus"
18544. .18650
repeat_region /note="MIR repeat: matches 137. .261 of consensus"
18799. .18820
repeat_region /note="11 copies 2 mer aa 100% conserved"
18835. .19184
misc_feature /note="THE1A repeat: matches 1. .354 of consensus"
19231. .19522
repeat_region /note="match: STS: Em:HS153YF5"
19297. .19346
repeat_region /note="25 copies 2 mer ca 100% conserved"
19426. .19564
repeat_region /note="MIR repeat: matches 6. .147 of consensus"
19624. .19916
repeat_region /note="AluX repeat: matches 1. .295 of consensus"
20005. .20204
repeat_region /note="MER20 repeat: matches 1. .218 of consensus"
20215. .20317
repeat_region /note="MER5A repeat: matches 4. .108 of consensus"
20410. .20541
repeat_region /note="MER5B repeat: matches 1. .120 of consensus"
20544. .20687
repeat_region /note="AluSg/x repeat: matches 154. .297 of consensus"
20689. .20814
repeat_region /note="FLAM A repeat: matches 14. .136 of consensus"
20816. .21218
repeat_region /note="Charlie4a repeat: matches 17. .427 of consensus"
21337. .21634
repeat_region /note="AluY repeat: matches 1. .297 of consensus"

Query Match 92.1%; Score 1387.2; DB 9; Length 128871;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
QY 2 CAGGTTGCGGATGACTCTTGTCAATTCGAACACTCATTCATGAGCAAGAGC 61
DB 71583 CAGGTTGCGGATGACTCTTGTCAATTCGAACACTCATTCATGAGCAAGAGA 71524
QY 62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGCATTTGGAACTTGTATGCTTGT 121
DB 71523 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGCATTTGGAACTTGTATGCTTGT 71464
QY 122 TTGTGAAAGAACCACTGACACATCACTGAGCTTCTTAAAGTTGCAAGAGTTAGAGG 181
DB 71463 TTGTGAAAGAACCACTGACACATCACTGAGCTTCTTAAAGTTGCAAGAGTTAGAGG 71404
QY 182 ACTATACACTTTCTTTTGAACCTTTATATAATAATTTGCTCTGG-TTTTGAAACCCAGG 240
DB 71403 ACTATACACTTTCTTTTGAACCTTTATATAATAATTTGCTCTGGTTTGTGAAACCCAGG 71344
QY 241 ACTGTTTGA-GGTGAGTGACAGGCTTAC-AGTGGCCTTTAATCCAACTCCAGAAATTGC 298
DB 71343 GCTGTTAGAGGGTGAGTGACAGCTTTAAGTGGCCTTTATCCAACTCCAGAAATTGC 71284
QY 299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358
DB 71283 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 71224
QY 359 CTCCTTAATGTACATGAT-GGGCAAGAGTATGGAGCTCTCTTGGCAGTCCGATGGAG 417
DB 71223 CTCCTTAATGTACATGATGGGCAAGTGTGGCAGCTCTCTTGGCAGTCCGATGGAG 71164
QY 418 ATGGA-GATGCTTGTCAATGAAAGGGCCNCTGTTGCAATTTCCGAGCTACACAAAGAA 476
DB 71163 ATGGAGGATGCTTGTCAATGAAAGGGCCGCTGTTGTTCCATTCGAGCTACACAAAGAA 71104
QY 477 AAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTTGGATTGCATTTCTTGCGCCAGAC 536

DB 71103 AAAAAATGT-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATTTCTGCGCCAGAC 71045
QY 537 CTTTCACATTCAGAAAGACCTTTAATAAACATGCTTTAATGCAACACCGGCTTACCTCTG 596
DB 71044 CTTTCACATTCAGAAAGACCTTTAATAAACATGCTTTAATGCAACACCGGCTTACCTCTG 70985
QY 597 TGAACAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 656
DB 70984 TGAACAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 70925
QY 657 GCGAAGCAACCTCCAGGAAAGAAATTCGAAGAAATGAATTTAGCTGTGAGGTATG 716
DB 70924 GCGAAGCAACCTCCAGGAAAGAAATTCGAAGAAATGAATTTAGCTGTGAGGTATG 70865
QY 717 TGGGCGAGCATTTAGAGTCGCTTTTGTGATTTGAGATCCACATGAGAAACACACAAAGATT 776
DB 70864 TGGGCGAGCATTTAGAGTCGCTTTTGTGATTTGAGATCCACATGAGAAACACACAAAGATT 70805
QY 777 TTTCACTTACGGGTTAACAATGTCGGAAAGAAATTCGAAGAGCTTGGTTCTTAAAAA 836
DB 70804 TTTCACTTACGGGTTAACAATGTCGGAAAGAAATTCGAAGAGCTTGGTTCTTAAAAA 70745
QY 837 TCACATGCGGACRCAATATGCGAAATCGGGGCCGAGAACCAACTGCAGCAAGGCTTGA 896
DB 70744 TCACATGCGGACRCAATATGCGAAATCGGGGCCGAGAACCAACTGCAGCAAGGCTTGA 70685
QY 897 GAGTAGTCAGCAACGATCAACGAGGTGTCAGAGTGTCACGCGCGCGAGAGCATCTCTC 956
DB 70684 GAGTAGTCAGCAACGATCAACGAGGTGTCAGAGTGTCACGCGCGCGAGAGCATCTCTC 70625
QY 957 TCCTTGCAAAATCTGCATGGTTTGTGGTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 1016
DB 70624 TCCTTGCAAAATCTGCATGGTTTGTGGTTCCTATTTCCAAATAAAGAAAGTCTAATTGA 70565
QY 1017 GCACGCAAGGTGACACACCAAAAAAATCTGCTTTCCGTACACAGCGCGCGAGAGCATCT 1076
DB 70564 GCACGCAAGGTGACACACCAAAAAAATCTGCTTTCCGTACACAGCGCGCGAGAGCATCT 70505
QY 1077 TCACAGAGGAGAAATGCGGTCTCGAGGAGGACCTTCTGCAAGTTGTTCAACTTGAGACC 1136
DB 70504 TCACAGAGGAGAAATGCGGTCTCGAGGAGGACCTTCTGCAAGTTGTTCAACTTGAGACC 70445
QY 1137 AAAATCTCACCTCGAAACGGGGAAGAGCTGTGAGATGCAATCCCTCAGCTCGATCCGTT 1196
DB 70444 AAAATCTCACCTCGAAACGGGGAAGAGCTGTGAGATGCAATCCCTCAGCTCGATCCGTT 70385
QY 1197 CACCACTTTCAGGCTTGGCACTGCTACAAAGAAAGTTGCCATTTGCCAAGAGT 1256
DB 70384 CACCACTTTCAGGCTTGGCACTGCTACAAAGAAAGTTGCCATTTGCCAAGAGT 70325
QY 1257 GAAGGAATTTGGGCAAGAGGAGGACACGACAAACGATTCGAGTTCCGAGAGGAGCT 1316
DB 70324 GAAGGAATTTGGGCAAGAGGAGGAGGACACGACAAACGATTCGAGTTCCGAGAGGAGCT 70265
QY 1317 TGGAGAAACAAATAAGAACCATTTGTGAGGCTCTCGAAGAAAGAGAGTGCAGAAACA 1376
DB 70264 TGGAGAAACAAATAAGGACAGTTGTGAGGCTCTCGAAGAAAGAGAGTGCAGAAACA 70205
QY 1377 CTCACAGCGGAGGCGCTTCCGTGGAACGCGATCCCAAGTTACCAAGTAGCAAGAGAA 1436
DB 70204 CTCACAGCGGAGGCGCTTCCGTGGAACGCGATCCCAAGTTACCAAGTAGCAAGAGAA 70145
QY 1437 GCCCACTCACTGCTCCGAGTGGGCAAGAGCTTTTCAGAACCTACCAAGCTGCTTGA 1496
DB 70144 GCCCACTCACTGCTCCGAGTGGGCAAGAGCTTTTCAGAACCTACCAAGCTGCTTGA 70085
QY 1497 CTCAGAGGCTCC 1507
DB 70084 CTCAGAGGCTCC 70074

AR157086	AR157086	5632 bp	DNA	linear	PAT 08-AUG-2001
LOCUS	Sequence 3 from patent US 6242590.				
DEFINITION	Antisense modulation of zinc finger protein-217 expression				
ACCESSION	AR157086.1	GI:15125790			
VERSION	Unknown.				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 5632)				
AUTHORS	Cowser, L.M.				
TITLE	Antisense modulation of zinc finger protein-217 expression				
JOURNAL	Patent: US 6242590-A 3 05-JUN-2001;				
FEATURES	Location/Qualifiers				
source	1..5632				
	/organism="unknown"				
	/mol_type="unassigned DNA"				
ORIGIN					
Query Match	88.7%; Score 1337.2; DB 6; Length 5632;				
Best Local Similarity	97.2%; Pred. No. 0;				
Matches 1420; Conservative	3; Mismatches 32; Indels 6; Gaps 6;				
QY	52 GACAAAGAGCACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCCCTG 111				
Db	1 GACAAAGAGCACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCCCTG 60				
QY	112 TATGCTTGTGTTGGAAAGAACAGTGCACCATCATCTAGCTTCCTAAAGTTGGAAG 171				
Db	61 TATGCTTGTGTTGGAAAGAACAGTGCACCATCATCTAGCTTCCTAAAGTTGGAAG 120				
QY	172 AAGTTAGAGGACTATACACTTCTTTTGAACCTTTTATATTAATTAATTTGCTCTGG-TTTT 230				
Db	121 AAGTTAGAGGACTATACACTTCTTTTGAACCTTTTATATTAATTAATTTGCTCTGGTTTT 180				
QY	231 GGAAACCCAGGACTGTTAGA-GGGTGAGTGACAGGCTTTAC-AGTGGCCCTTAATCCAACTC 288				
Db	181 GGAACCCAGGCTGTTAGGGGCTGAGTGACAACTTTACAGTGGCCCTTAATCCAACTC 240				
QY	289 CAGAAATTTGCCCAACGGAACTTTGAGATTTATATGCAATCGAAAGTGACAGGAAACATGCC 348				
Db	241 CAGAAATTTGCCCAACGGAACTTTGAGATTTATATGCAATCGAAAGTGACAGGAAACATGCC 300				
QY	349 AACTCAATCCCTCTTAATGTACATGAT-GGCCAAGAGTGATTTGCGAGCTCTTTGCCAG 407				
Db	301 AACTCAATCCCTCTTAATGTACATGATGGCCAGAGTGATTTGCGAGCTCTTTGCCAG 360				
QY	408 TCCGATGGAGATGGA-GATGCCCTTGTCAATGAAGGGCCNCCTGTTGTCAATTTCCGAGCT 466				
Db	361 TCCGATGGAGATGGAGGATGCCCTTGTCAATGAAGGGCCNCCTGTTGTCAATTTCCGAGC 420				
QY	467 ACACAAAGAAATAATGCAATCCGAATCGAGGGGAATATGCCCTTTGGATTCGATGTTCT 526				
Db	421 TACACAAAGAAATAATGCAATCCGAATCGAGGGGAATATGCCCTTTGGATTCGATGTTCT 479				
QY	527 GCAGCCAGACCTTCACACATTCAGAGACCTTAATAAATCTCTTAATGCAACACCGGC 586				
Db	480 GCAGCCAGACCTTCACACATTCAGAGACCTTAATAAATCTCTTAATGCAACACCGGC 539				
QY	587 CTACCCCTCTGTGAACACGAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATA 646				
Db	540 CTACCCCTCTGTGAACACGAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATA 599				
QY	647 AAAGTCAAGTGGGAACAGAACTCCCAAGGAAAGAAATTTGCAAGAAATGAATTTAGCT 706				
Db	600 AAAGTCAAGTGGGAACAGAACTCCCAAGGAAAGAAATTTGCAAGGAAATGAATTTAGCT 659				
QY	707 GTGAGGTATGGGAGACATTTAGAGTGCCTTTTGTGATTTGAGATCCACATGAGAACAC 766				
Db	660 GTGAGGTATGGGAGACATTTAGAGTGCCTTTTGTGATTTGAGATCCACATGAGAACAC 719				
QY	767 ACAAGATTCTTTCACTTACGGGTGTAAACATGTGCGGAAGAGATTCAAGGAGCCCTTGT 826				
Db	767 ACAAGATTCTTTCACTTACGGGTGTAAACATGTGCGGAAGAGATTCAAGGAGCCCTTGT 826				
Db	720 ACAAGATTCTTTCACTTACGGGTGTAAACATGTGCGGAAGAGATTCAAGGAGCCCTTGT 779				
QY	827 TTCCTTAAAAATCAGATCGGACRCATATATGCAATCGGGGCCAGAGCAACTGCAGC 886				
Db	780 TTCCTTAAAAATCAGATCGGACRCATATATGCAATCGGGGCCAGAGCAACTGCAGC 839				
QY	887 AAGGCTTGGAGATAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGCAACGCGGCCGAGA 946				
Db	840 AAGGCTTGGAGATAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGCAACGCGGCCGAGA 899				
QY	947 GCATCTCTCTCTCTTTCGCAAAATTCGATGTTGTTGCTTCTTATTTCAATAAAGAAA 1006				
Db	900 GCATCTCTCTCTCTTTCGCAAAATTCGATGTTGTTGCTTCTTATTTCAATAAAGAAA 959				
QY	1007 GTCTAATTTGACACCGGCAAGGTGCACACCAAAAACCTGCTTTCGGTACCAGCGCGC 1066				
Db	960 GTCTAATTTGACACCGGCAAGGTGCACACCAAAAACCTGCTTTCGGTACCAGCGCGC 1019				
QY	1067 AGACAGACTCTCCCAAGGAGGATGCCGCTCTCCAGGGAGGACTTCTCTGCAAGTTGTTCA 1126				
Db	1020 AGACAGACTCTCCCAAGGAGGATGCCGCTCTCCAGGGAGGACTTCTCTGCAAGTTGTTCA 1079				
QY	1127 ACTTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCTGTTCAGATGCATCCCTCAGC 1186				
Db	1080 ACTTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCTGTTCAGATGCATCCCTCAGC 1139				
QY	1187 TCGATCCGTTTACCACCTTCCAGGCTTGGCACTGGCTACCAAGGAAGTTGCCATTT 1246				
Db	1140 TCGATCCGTTTACCACCTTCCAGGCTTGGCACTGGCTACCAAGGAAGTTGCCATTT 1199				
QY	1247 GCCAAGAGTGAAGAAATTGGGGCAAGAGGAGCACCGACAAACGACGATTCGAGTTCG 1306				
Db	1200 GCCAAGAGTGAAGAAATTGGGGCAAGAGGAGCACCGACAAACGACGATTCGAGTTCG 1259				
QY	1307 AGAAGGAGCTTTGGAGAAACAAATTAAGAACCAATTGTGAGGCTCTTCGCAAGAGAAAGA 1366				
Db	1260 AGAAGGAGCTTTGGAGAAACAAATTAAGAACCAATTGTGAGGCTCTTCGCAAGAGAAAGA 1319				
QY	1367 AGTGCMAACATCCCAACGGCGAGGCGCTCCGTGGACGGGATCCCAAGTTACCCAGTA 1426				
Db	1320 AGTGCMAACATCCCAACGGCGAGGCGCTCCGTGGACGGGATCCCAAGTTACCCAGTA 1379				
QY	1427 GCAAGGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACTTACCACAGC 1486				
Db	1380 GCAAGGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACTTACCACAGC 1439				
QY	1487 TGGTCTTGCATCCAGGCTCC 1507				
Db	1440 TGGTCTTGCATCCAGGCTCC 1460				
RESULT 8					
AF041259	5632 bp	mRNA	linear	PRI 22-JUL-1998	
LOCUS	Homo sapiens breast cancer putative transcription factor (ZBC1)				
DEFINITION	mRNA, complete cds.				
ACCESSION	AF041259				
VERSION	AF041259.1	GI:3335396			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
	1 (bases 1 to 5632)				
	Collins, C., Rommens, J.M., Kowbel, D., Godfrey, T., Tanner, M.,				
	Hwang, S.-I., Polikoff, D., Nonet, G., Cochran, J., Myambo, K.,				
	Jay, K.E., Froula, J., Cloutier, T., Kuo, W.-L., Yaswen, P., Dairkee, S.,				
	Giovannola, J., Hutchinson, G.B., Isola, J., Kallioniemi, O.-P.,				
	Palazzolo, M., Martin, C., Ericsson, C., Pinkel, D., Albertson, D.,				
	Li, W.-B. and Gray, J.W.				
TITLE	Positional cloning of ZNF217 and NAB1: genes amplified at 20q13.2				
JOURNAL	and overexpressed in breast carcinoma				
	Proc. Natl. Acad. Sci. U.S.A. 95 (15), 8703-8708 (1998)				

MEDLINE	98337979	Query Match	88.7%;	Score 1337.2;	DB 9;	Length 5632;
PUBMED	9671742	Best Local Similarity	97.2%;	Pred. No. 0;		
REFERENCE	2 (bases 1 to 5632)	Matches 1420;	Conservative	3;	Mismatches 32;	Indels 6;
AUTHORS	Hwang,S.-I., Polikoff,D., Nonet,G., Cochran,J., Myambo,K., Jay,K.E., Froula,J., Cloutier,T., Kuo,W.-L., Yaswen,P., Dairkee,S., Giovanola,J., Hutchinson,G.B., Isola,J., Kallioniemi,O.-P., Palazzolo,M., Martin,C., Ericsson,C., Pinkel,D., Albertson,D., Li,W.-B. and Gray,J.W.					
TITLE	Direct Submission					
JOURNAL	Submitted (07-JAN-1998) Life Sciences, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA					
FEATURES	Location/Qualifiers					
source	1..5632					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/chromosome="20"					
	/map="20q13.2-q13.3"					
gene	1..5632					
CDS	/gene="ZABC1"					
	272..3418					
	/gene="ZABC1"					
	/note="Kruppel-like"					
	/codon_start=1					
	/product="breast cancer putative transcription factor"					
	/protein_id="AAC39895.1"					
	/db_xref="GI:3335397"					
	/translation="MQSKVTGNMPTQSLMYMDGPEVIGSSISGSPMEMEDALSMSKGTAVVPPRATQENKVIQIEGYMPLDCMFCQSOTFTHSEDLNKHVLMQHRPTLCEPAVLVEAEYLSPLDKQVREPPKRNKTHNGKSGARSKLQGGLESSPATINEVQVHAESISSPYKICCGRPFKBPFLMPPKTHNGKSGARSKLQGGLESSPATINEVQVHAESISSPYKICMYCGFLFPNKESLIEHRKVTHTKTAFGTSSAQTQDQGGKSSREDFLQLNLRPKSHPTGKKPVCICI POLDPFTTFQAWOLATKGKVAICQEVKSGOEGSDNDSSSEKELGENTKSGCAGLQEKKEKSHGEAPSVADADPKLPSSKRPKTHCSECKAFRTYHOLVLSHVNKDRQAGAESPTMSVDGROPGTCSPLDAAPLDENGAVDRGEGSEDSERGLPEGHLDNDGGRKIRHLTSRESCYCGKFRSNIYLHRTHTGKPKYKCFEYAAAQKTSRHLRHHKKEQTDVAAEVKNDQKQDTEALLTADSAQTNKLRPFDDGKLDVTGSPKQLKEMPSVFQNLGSAVLSPAHKQTDQPHKNAADSDKVNKNPTPAYLDLKKRSVETQANNLIRCTKADVTTPDGGSTTHNLEVPKEQKQETETAADCRVPSVDCHEKPLNLSVGLAHNCPAISLSKSLIPSIITCPCTKTFYFVPLVMHQRLEHKYNPVDVHKCRNKSLLRSRTPGCPALLGKDVPPLLSFCFKPKPSAPPAQSKLSPSAKGQSPPGPKATKLPVAGSDSLAPSLNKLKSHRPQONVGVGAATRQOQSEMFPKTSVSPADPTKRPTKGLKPLVAVDQPCANYRRGYDLPKYHVRGITSLLPQDCVYPSQALPPKPRFLSSSEVDSPNVLTVOKPYGSGGLYTCVPAGSPASSSTLEGKRPVSYQHLNSMAQKRNYENFIGNAHYRPNDKKT"					
misc_feature	470..535					
	/gene="ZABC1"					
	/note="zinc finger 1"					
misc_feature	659..721					
	/gene="ZABC1"					
	/note="zinc finger 2"					
misc_feature	744..805					
	/gene="ZABC1"					
	/note="zinc finger 3"					
misc_feature	923..985					
	/gene="ZABC1"					
	/note="zinc finger 4"					
misc_feature	1400..1463					
	/gene="ZABC1"					
	/note="zinc finger 5"					
misc_feature	1688..1750					
	/gene="ZABC1"					
	/note="zinc finger 6"					
misc_feature	1772..1834					
	/gene="ZABC1"					
	/note="zinc finger 7"					
misc_feature	2405..2470					
	/gene="ZABC1"					
	/note="zinc finger 8"					
ORIGIN						

Db	1020	AGACAGACTCTCCACAGGAGGAATGCCGTCTCCAGGGAGGACTTCTCGAGTTGTTCA	1079	
Qy	1127	ACTTGAGACCAAAATCTCAACCTGAAACGGGGAAGACGCTGTGATGCAATCCCTCAGC	1186	
Db	1080	ACTTGAGACCAAAATCTCAACCTGAAACGGGGAAGACGCTGTGATGCAATCCCTCAGC	1139	
Qy	1187	TCGATCCGTTACACACCTTCCAGGCTTGGCACTGCTGATCCAAAGAAAGTTGCAATTT	1246	
Db	1140	TCGATCCGTTACACACCTTCCAGGCTTGGCACTGCTGATCCAAAGAAAGTTGCAATTT	1199	
Qy	1247	GCCAAAGAGTGAAGAAATTTGGGGAAGAGGAGCACCGACAAACGACGATTCGAGTTCCG	1306	
Db	1200	GCCAAAGAGTGAAGAAATTTGGGGAAGAGGAGCACCGACAAACGACGATTCGAGTTCCG	1259	
Qy	1307	AGAAGGAGCTTGGAAGAAACAAATAAGAACCAATTTGTGAGGCTCTTCGAAAGAAAGAGA	1366	
Db	1260	AGAAGGAGCTTGGAAGAAACAAATAAGAGGAGTGTGTGACAGGCTCTTCGAAAGAAAGAGA	1319	
Qy	1367	AGTGCAACACTCCACCGGAGGCGCTCCGTCGGAGCGCGATCCCAAGTTACCCAGTA	1426	
Db	1320	AGTGCAACACTCCACCGGAGGCGCTCCGTCGGAGCGCGATCCCAAGTTACCCAGTA	1379	
Qy	1427	GCAAGGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGAGCTTTCAGAACCTACCAACGAGC	1486	
Db	1380	GCAAGGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGAGCTTTCAGAACCTACCAACGAGC	1439	
Qy	1487	TGGTCTTGACATCCAGGGTCC	1507	
Db	1440	TGGTCTTGACATCCAGGGTCC	1460	
RESULT 9				
LOCUS CQ726295				
DEFINITION Sequence 12229 from Patent WO02068579.				
ACCESSION CQ726295				
VERSION CQ726295.1 GI:42288937				
KEYWORDS				
SOURCE Homo sapiens (human)				
ORGANISM Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
1				
REFERENCE Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.				
AUTHORS Kits, such as nucleic acid arrays, comprising a majority of				
TITLE humanexons or transcripts, for detecting expression and other uses				
thereof				
JOURNAL Patent: WO 02068579-A 12229 06-SEP-2002;				
PE Corporation (NY) (US)				
FEATURES				
source				
1..5633				
Location/Qualifiers				
/organism="Homo sapiens"				
/mol_type="unassigned DNA"				
/db_xref="taxon:9606"				
ORIGIN				
Query Match 88.7%; Score 1337.2; DB 6; Length 5633;				
Best Local Similarity 97.2%; Pred. No. 0;				
Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;				
Qy	52	GACAAAGAGCACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCAATGGGAACCCGTG	111	
Db	1	GACAAAGAGCACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCAATGGGAACCCGTG	60	
Qy	112	TATGCTTGTGTTGTGGAAGAACCAAGTACACCATCACTGAGCTTCTTAAAGTTTCGAAG	171	
Db	61	TATGCTTGTGTTGTGGAAGAACCAAGTACACCATCACTGAGCTTCTTAAAGTTTCGAAG	120	
Qy	172	AAGTTAGAGCACTATACACTTCTTTTGAACCTTTTATATAATATTTGCTCTGG-TTTT	230	
Db	121	AAGTTAGAGCACTATACACTTCTTTTGAACCTTTTATATAATATTTGCTCTGGTTTTT	180	

Qy	231	GGAAACCCAGGACTGTTTAGA-GGGTGAGTGACAGGCTCTTAC-AGTGGCCTTAATCCAACTC	288	
Db	181	GGAAACCCAGGCTGTTAGAGGGGTGAGTGACAGTCTTTACAAGTGGCCTTATTTCCAACTC	240	
Qy	289	CAGAAATTTGCCCAACGGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC	348	
Db	241	CAGAAATTTGCCCAACGGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC	300	
Qy	349	AATCAATCCCTCTTAATGTACATGGAT-GGCCAAGAGTGATTTGGCAGCTCTCTTGGCAG	407	
Db	301	AATCAATCCCTCTTAATGTACATGGATGGGCGCAAGATGATTTGGCAGCTCTCTTGGCAG	360	
Qy	408	TCCGATGAGATGGA-GATGCTTGTCAATGAAAGGGCCNCCTGTGTCAATTCGAGCT	466	
Db	361	TCCGATGAGATGAGGATGCTTGTCAATGAAAGGGACCGCTGTGTTCATTCGAGCT	420	
Qy	467	ACAAAGAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT	526	
Db	421	TACACAAGAAAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT	479	
Qy	527	GCAGCAGACCTTCACACATTCAGAGACCTTATAACAATGTCTTAATGCACACCGC	586	
Db	480	GCAGCAGACCTTCACACATTCAGAGACCTTATAACAATGTCTTAATGCACACCGC	539	
Qy	587	CTACCTCTGTGAACACGACGCTTCTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATA	646	
Db	540	CTACCTCTGTGAACACGACGCTTCTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATA	599	
Qy	647	AAAGTCAAAGTGCAGAACAGAACCTCCCAAGGAAAAAGATTGCAAGGAAAAATGAATTAGCT	706	
Db	600	AAAGTCAAAGTGCAGAACAGAACCTCCCAAGGAAAAAGATTGCAAGGAAAAATGAATTAGCT	659	
Qy	707	GTGAGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACAC	766	
Db	660	GTGAGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACAC	719	
Qy	767	ACAAAGATTCTTTTCACTTACCGGTGTAAACATGTGCGGAAGAAAGATTCAAGAGCCTTGGT	826	
Db	720	ACAAAGATTCTTTTCACTTACCGGTGTAAACATGTGCGGAAGAAAGATTCAAGAGCCTTGGT	779	
Qy	827	TTCTTAAAAATCACATGCGGACRCATATATGCAAAATCGGGGGCCAGAGCAAACTGCAGC	886	
Db	780	TTCTTAAAAATCACATGCGGACACATATATGCAAAATCGGGGGCCAGAGCAAACTGCAGC	839	
Qy	887	BAAGCTTGGAGTAGTCCAGCAGCATCAAGAGTGTGTCAGAGTGCACCGGCGCGGAGA	946	
Db	840	BAAGCTTGGAGTAGTCCAGCAGCATCAAGAGTGTGTCAGAGTGCACCGGCGCGGAGA	899	
Qy	947	GCATCTCTCTCTTGCAGAAATCTGCATGTTGTTGGCTTCTTATTTCCAAATAAGAAA	1006	
Db	900	GCATCTCTCTCTTGCAGAAATCTGCATGTTGTTGGCTTCTTATTTCCAAATAAGAAA	959	
Qy	1007	GTCATAATTGAGACCGCAAGGTGCACACCAAAAAAATCTGCTTTTCGGTACCAGCAGCGGC	1066	
Db	960	GTCATAATTGAGACCGCAAGGTGCACACCAAAAAAATCTGCTTTTCGGTACCAGCAGCGGC	1019	
Qy	1067	AGACAGACTCTCCAAAGGAGGATGCGCTCTCGAGGGAGCACTTCTCGAGTTGTTCA	1126	
Db	1020	AGACAGACTCTCCAAAGGAGGATGCGCTCTCGAGGGAGCACTTCTCGAGTTGTTCA	1079	
Qy	1127	ACTTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCCTGTTCAGATGCATCCCTCAGC	1186	
Db	1080	ACTTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCCTGTTCAGATGCATCCCTCAGC	1139	
Qy	1187	TCGATCCGTTCCACCTTCCAGGCTTGGCAKCTGGCTTACAAAGGAAGTTGCAATTT	1246	
Db	1140	TCGATCCGTTCCACCTTCCAGGCTTGGCAKCTGGCTTACCAAGGAAGTTGCAATTT	1199	
Qy	1247	GCCAAAGAGTGAAGAAATTTGGGGAAGAGGAGCACCGACAAACGACGATTCGAGTTCCG	1306	
Db	1200	GCCAAAGAGTGAAGAAATTTGGGGAAGAGGAGCACCGACAAACGACGATTCGAGTTCCG	1259	
Qy	1307	AGAAGGAGCTTTGGGAAAAACAAATAAGAACCTTGTGCGAGGCTCTCTCGCAAGAGAAAGAGA	1366	

```

|||||
1260 AGAAGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCAGGCTCTCCGAGGAAAGAGA 1319
|||||
1367 AGTGCAAAACATCCCAACGGCGAAGCCCTCCGTGGACGGATCCCAAGTTACCCAGTA 1426
|||||
1320 AGTGCAAAACATCCCAACGGCGAAGCCCTCCGTGGACGGATCCCAAGTTACCCAGTA 1379
|||||
1427 GCAAGGAGAGCCCACTCACTGCTCGAGTGGGCAAAAGCTTTCAGAACTTACCACGAC 1486
|||||
1380 GCAAGGAGAGCCCACTCACTGCTCGAGTGGGCAAAAGCTTTCAGAACTTACCACGAC 1439
|||||
1487 TGGTCTTGCACCTCCAGGCTCC 1507
|||||
1440 TGGTCTTGCACCTCCAGGCTCC 1460
|||||

RESULT 10
BD085734
LOCUS
DEFINITION
BD085734 3186 bp DNA linear PAT 27-AUG-2002
ACCESSION
BD085734
VERSION
BD085734.1 GI:226311344
KEYWORDS
JP 2001524802-A/10.
SOURCE
synthetic construct
ORGANISM
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 3186)
AUTHORS
Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and
Romens,J.
Genes from the 20q13 amplicon and their uses
Patent: JP 2001524802-A 10 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS Artificial Sequence
PN JP 2001524802-A/10
PD 04-DEC-2001
PF 15-JUN-1997 JP 1998506264
PR 15-JUL-1996 US 08/660395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
PI JOHANNA ROMENS
PC C12N15/11,C12Q1/68,A61K48/00
CC Description of Artificial Sequence:ZABCI Open Reading Frame FH
Key
Location/Qualifiers
FT source 1..3186
/organism='Artificial Sequence'.
FEATURES
source
1..3186
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'

ORIGIN
Query Match 73.6%; Score 1108.6; DB 6; Length 3186;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1156; Conservative 3; Mismatches 28; Indels 3; Gaps 3;

QY 320 ATGCAATCAAAGTGCACGAAACATGCAACTCAATCCCTCTTAATGTACATGATGG- 378
|||||
1 ATGCAATCAAAGTGCACGAAACATGCAACTCAATCCCTCTTAATGTACATGATGGG 60
|||||
QY 379 CCAAGAGTATGGGAGCTCTCTTGGCAGTCCGATGGAGATGGA-GATGCCCTTGTCAATG 437
|||||
DB 61 CCAGAAGTATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGTGCCTTGTCAATG 120
|||||
QY 438 AAAGGCCNCCTGTTGTCAATTCGAGCTACACAAGAAAAAATGTCAATCCGAATCGA 497
|||||
DB 121 AAAGGACCGCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTG-ATCCAAATCGA 179
|||||
QY 498 GGGGATATGCCCTGGATGTGATGTTCTGAGCCAGACCTTCACACATTCAGAGACCT 557
|||||
DB 180 GGGGTATATGCCCTGGATGTGATGTTCTGAGCCAGACCTTCACACATTCAGAGACCT 239
|||||

```

```

QY 558 TAAATAAACAATGCTCTTAATGCAACACACGGGCTACCTCTGTGTGAAACACGACGATCTTCGGGT 617
DB 240 TAAATAAACAATGCTCTTAATGCAACACACGGGCTACCTCTGTGTGAAACACGACGATCTTCGGGT 299
|||||
QY 618 TGAACGACGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCAGGCAACAGAACCTCCCAAGGA 677
DB 300 TGAACGACGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCAGGCAACAGAACCTCCCAAGGA 359
|||||
QY 678 AAAGAATTCGAAGGAAAAATGAAATTTAGCTGTGAGGTATGTGGGCGAGACATTTAGAGTCGC 737
DB 360 AAAGAATTCGAAGGAAAAATGAAATTTAGCTGTGAGGTATGTGGGCGAGACATTTAGAGTCGC 419
|||||
QY 738 TTTTGAATGTTGAGATCCACATGAGAACACACAAAAGATCTTTTCACTTACGGGTGTAAACAT 797
DB 420 TTTTGAATGTTGAGATCCACATGAGAACACACAAAAGATCTTTTCACTTACGGGTGTAAACAT 479
|||||
QY 798 GTGCGGAAAGATTCAAAGGAGCCTTTGTTTCTTAAATAATCACATGCGGACATCAATG 857
DB 480 GTGCGGAAAGATTCAAAGGAGCCTTTGTTTCTTAAATAATCACATGCGGACATCAATG 539
|||||
QY 858 CAAATCGGGGGCCAGAGCAAACTGACGCAAGGCTTTGGAGAGTAGTCCAGCAACGATCAA 917
DB 540 CAAATCGGGGGCCAGAGCAAACTGACGCAAGGCTTTGGAGAGTAGTCCAGCAACGATCAA 599
|||||
QY 918 CGAGGTGCTCAGGTGACGCGGCGGAGAGCATCTCTCTCTTGCATAAATCTGCATGTT 977
DB 600 CGAGGTGCTCAGGTGACGCGGCGGAGAGCATCTCTCTCTTGCATAAATCTGCATGTT 659
|||||
QY 978 TTGTGCTCTCTTATTTCCAAATAAAGAAAGTCTTAATTGAGCAGCCGAGGTGCACACCAA 1037
DB 660 TTGTGCTCTCTTATTTCCAAATAAAGAAAGTCTTAATTGAGCAGCCGAGGTGCACACCAA 719
|||||
QY 1038 AAAAACTGCTTTGCGGTACCCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCGGTC 1097
DB 720 AAAAACTGCTTTGCGGTACCCAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCGGTC 779
|||||
QY 1098 CTCGAGGAGGAGTCTCTGAGTGTTCACCTTGAAGACCAAAATCTCACCTTGAACGCG 1157
DB 780 CTCGAGGAGGAGTCTCTGAGTGTTCACCTTGAAGACCAAAATCTCACCTTGAACGCG 839
|||||
QY 1158 GAAGAGCTGTCAGATGCATCCCTCAGCTCGATCGTTCACCACTTCCAGCTTCCAGGCTTGGA 1217
DB 840 GAAGAGCTGTCAGATGCATCCCTCAGCTCGATCGTTCACCACTTCCAGCTTCCAGGCTTGGA 899
|||||
QY 1218 KCTGGCTACCAAGGAAGTGGCCATTTGCCAAGAGTGAAGGAATTTGGGCAAGAGG 1277
DB 900 GCTGGCTACCAAGGAAGTGGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAGG 959
|||||
QY 1278 GAGCACCGCAACACGACGATTCGAGTTCGAGAGGAGGCTTTGGAGAAACAAAATAAGAACCA 1337
DB 960 GAGCACCGCAACACGACGATTCGAGTTCGAGAGGAGGCTTTGGAGAAACAAAATAAGGGCAG 1019
|||||
QY 1338 TTGTGAGGCTCTCGCAAGAGAAAGAGAGTGGCAAACTCCCGGCGGAGGCGGCTC 1397
DB 1020 TTGTGAGGCTCTCGCAAGAGAAAGAGAGTGGCAAACTCCCGGCGGAGGCGGCTC 1079
|||||
QY 1398 CGTGACGCGGATCCCAAGTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTG 1457
DB 1080 CGTGACGCGGATCCCAAGTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTG 1139
|||||
QY 1458 CGGCAAGCTTTAGAACCTTACCACGAGTGGTCTTGCACTCCAGGGTCC 1507
DB 1140 CGGCAAGCTTTAGAACCTTACCACGAGTGGTCTTGCACTCCAGGGTCC 1189
|||||

RESULT 11
AC118373
LOCUS
DEFINITION
Rattus norvegicus clone CH230-35619, *** SEQUENCING IN PROGRESS
***, 6 unordered pieces.
ACCESSION
AC118373
VERSION
AC118373.3 GI:23194641
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

```


Db 139495 GCTTCTGGAGGTTTGAGTGGTGATATCTCAAGTGTCTTATTTAACTTCAGAGATCC 139554
QY 298 CCCAAGCGAATTTGAGATTATATGCA-----ATCGAAAGTGACAGGAACATGCCAACT 352
Db 139555 TTAGCAGACAGCTGGATCGTCCATCACTCAGACCTGAGTGGCCGCGAGCATGCCGAG 139614
QY 353 CAATCCCTCTTAATGTACATGGA-TGGCCAAAGATGATTGGCAGCTCTCTTTCGAGTCCG 411
Db 139615 CAATCCCTCTCTGTGTACATGGAACGCGCGGAGTCTCTTGGCAGTTCCTTAGTATCCAG 139674
QY 412 ATGGAGATGGA-GATGCTTGTCAATGAAGGCGCCNCTGTGTCAATTCGAGCTTACAC 470
Db 139675 ATGGAGGTGATGATGCTGTGCCCATGAAGGACCGGCTGTGTGCTCCCTTC-CGAGCTGC 139733
QY 471 AAAGAAAAAATGTCAATCGAATCGAGGGGAATATGCCCTTGGATTGGATGTCATGTTCTGCGAG 530
Db 139734 CGAGGAGAGAGCTTGGCCACAGCTCAGGGACACATGCCCCCTGGAGCTGCATGTTCTGCGAG 139793
QY 531 CGAGACCTTCACACATTTCAAGAGACCTTAAATAAATGTCCTTAATGCAACACCGGCTTAC 590
Db 139794 CCAGGTCTTCTCTCAGCGGAAGATCTCAATCAGCATGTGCTCTCTGACGACCGGCCAC 139853
QY 591 CTTCTGTGAACAGCAGTCTTCTCGGCTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAG 650
Db 139854 CTTCTGTGAGCGGCTGTGCTGTGTGAGGCGGAGTATTTAAGCCCCCTCGACAAAGG 139913
QY 651 TCAAGTGCAGAACAGACCTCCCAAGGAAAGAAATTCGAAGGAAA-----ATGAATTTAGCTG 707
Db 139914 TCAGGTGCCACCGAACCACCAAGGAGAGAGCAGCAGCAAGAAATGGAAGAGCTGAGCTG 139973
QY 708 TGAGGTATGTGGGACAGACATTTAGAGTCGTTTTGATGTTGATGATCCATGAGAACACA 767
Db 139974 TGATGTATCGGACAGACATATCCAGTGGCTTTTGTATGTTGAGGACCATGAAGAAACA 140033
QY 768 CAAAGATCTTCTCAGTTCAGGCTGTAAATGTCGGAAGAAATTCAGAGGCTTGGTT 827
Db 140034 TAAGGACTCTTTCAGCTATGAGGTGACGATGTCGCGGAGGAGATTCAGAGGCGCGTGT 140093
QY 828 TCTTAAAAATCATATCGGACRCAATATGGCAAAATCGGCGGCGCAGAAAGCAAACTGCAGCA 887
Db 140094 CTGGAAGACACATGCGGACACACATATGCGAAGTCTGSCACCAAGAGCAAGCTGCAGCA 140153
QY 888 AGGCTTGGAGAGTAGTCCAGCAACGATCAACAGAGTGTCTCAGGTGCAACGCGGCGAGAG 947
Db 140154 GGGCAGCGAG---AGTCCAGTCAACATCAACGATGTGTGTCAGGACATGCTCCCTGGGAA 140210
QY 948 CATCTCTCTCTTGCAGAAATCTGATGCTTGTGGCTTCTTATTTCCAAATGAAGAAAG 1007
Db 140211 CATCTCCACTCCCTACAAGATCTGCATGCTGTGTGGCTTCTCTTCCCAATTAAGCAGAG 140270
QY 1008 TCTAATTGAGCACGCAAGGTGCACACCAAAAAAACTGCTTTCGTTACCAAGCAGCGCGCA 1067
Db 140271 CTTATTGAGCACAGAGGTTTCATGCCAAGAAATGTTCTTAGTGCAGCAGCAGCAGC 140330
QY 1068 GACAGATCTCCACAGGAGGAATCGCTCTCAGGAGGAGACTTCTCTGAGTGTGTCAA 1127
Db 140331 CCTGTAGTCCCAAGAGAGAACCCACATCCCGAGGAGAGAGTCTCTCAGTGTGTTGAA 140390
QY 1128 CTTGAGACCAAAATCTCACCTTGAACGGGAGAGAGCTGTTCAGATGATCCCTCAGCT 1187
Db 140391 CTTGAGACCAAGATCAACTCAGAGAGATGACAGTGAAGGCCCATGACCTCTATACCTCAGCT 140450
QY 1188 CGATCCGTTTCCACCTCTCAGGCTTGGCAKCTGGCTTACCAAGGAAGTGTGCCATTG 1247
Db 140451 TGACCAATTCACCACTTACAGGATGAGGATGGTGGCTTACCAAGGAAGTGTGCCATTG 140510
QY 1248 CCAAGAA---GTGAAGGAATTTGGGCAAGAGGAGCAGCAGCAACAGCATGTCAGATTTC 1304
Db 140511 CCAAGAAAGAGGTGAAGAGTTCAGGCCCAAGAGGAGCAGCAGCAACAGTATTCGTGTTTC 140570
QY 1305 CGAGA---AGGAGCTTGGAGAA----- 1323
Db 140571 AGAGAAAGAGGAACCTTGGAGAAATATATGGTGTGGGGTGAAGGTGGAAGGTTCTGGAAAGTC 140630

QY 1324 ----ACAAATAAGAACCATTTGTGCGGCTCTCGCAAGAGAAAGAGAGTGCAAACACTC 1379
Db 140631 CAAGACAAATAAGACAGTGTGTTCAAGTCTCTCCCAAGACAAAGATAAGCTAGACATGG 140690
QY 1380 CCACGCGAGAGCGCTTCCGTGACCGGATCCCAAGTTTACCAGTAGCAAGGAGAGCC 1439
Db 140691 TTAACAGTGAAGTGCCTTTCAGGGGACGCTGACTCCAAAGTTTCCCGGAGCAAGAGAGCC 140750
QY 1440 CACTCACTCTCCGAGTGGGCAAGCTTTTCAAGACCTACCACAGCTGGTCTTTCGACTC 1499
Db 140751 CACACACTGTTCCGAGTGCAGTAAAGCGTTTCAAGACATACCACAGCTGGTCTTCGACTC 140810
QY 1500 CAGGCTCC 1507
Db 140811 GAGGCTGC 140818

RESULT 12
AC099079
LOCUS
DEFINITION
AC099079 246332 bp DNA linear HTG 10-MAY-2003
Rattus norvegicus clone CH230-99J9, *** SEQUENCING IN PROGRESS ***,
6 unordered pieces.
AC099079
HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Rattus.
REFERENCE
1 (bases 1 to 246332)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayodeji,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,H., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,O., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Naif,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpunaagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,
Plopper,F., Poinexter,A., Popovic,D., Primus,E., Fu,L.-L.,
Puazo,M., Quirroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Stemle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,

Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

Direct Submission

TITLE
JOURNAL
REFERENCE

2 (bases 1 to 246332)

Worley, K.C.

Direct Submission

TITLE
JOURNAL

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 246332)

Direct Submission

TITLE
JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GPZC

Center clone name: CH230-99J9

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 209040 bases at least Q40

Consensus quality: 213312 bases at least Q30

Consensus quality: 216534 bases at least Q20

Estimated insert size: 219845; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 14481: contig of 14481 bp in length
* 14482 14581: gap of unknown length
* 14582 41011: contig of 26430 bp in length
* 41012 41111: gap of unknown length
* 41112 239819: contig of 198708 bp in length
* 239820 239919: gap of unknown length
* 239920 241122: contig of 1203 bp in length
* 241123 241222: gap of unknown length
* 241223 242316: contig of 1094 bp in length
* 242317 242416: gap of unknown length
* 242417 246332: contig of 3916 bp in length.

Location/Qualifiers

FEATURES

source
1. 246332
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-99J9"
1. 4672
/note="wgs_contig"
misc_feature
136991..139129
/note="wgs_contig"
misc_feature
174114..175268
/note="wgs_contig"

ORIGIN

Query Match 40.8%; Score 614.2; DB 2; Length 246332;
Best Local Similarity 68.8%; Pred. No. 4.2e-177;
Matches 1079; Conservative 3; Mismatches 397; Indels 89; Gaps 14;
QY 2 CAGGTTGCTGGGATTGACTTCTTGCTCAATTGAACACTCAATTCAATGGAGACAAGAGC 61
DB 50464 CACATTGCGGGATCAACGCTCTTGCTCAACTGAGCCACTCACCAGTGGAGATGACAGA 50523
QY 62 ACTAATGCTTTG-TGCTGATTCAATTTGAATCGAGGCAATGGGAACCTCTGATGCCCTTG 120
DB 50524 ACTGTTGCCCTCATGCTGACTCACTGCTGAATCACAGCCAGGCGAGCTGGATGCCCTTG 50583
QY 121 TTTGTGGAAGAACGAGTGACACCATCACTGAGTCTCTTAAATATTTTAAAGTTTGAAGTTAGAG 180
DB 50584 TTTGTGGAAGAACCA---CCTCTATCACTGAGTCTCCCCAAGCTC-----TG 50629
QY 181 GACTATACACTTCTTTTGAACCTTTTATAATAAATATTTGCTCTGTTTGGAAACCCAGG 240
DB 50630 GAATATACATCCCTCTGAACCTT-----AGAGTTGCTCTAATTTTGACGCCAGG 50680
QY 241 ACTGTTAGAGGTGAGTGACAGGCTTAC---AGTGGCTTAAATCCAACCTCAGAAATG 297
DB 50681 GCTTCTGGAGGTTTGAGTGGTGATATCTCAAGTGTCTTATTCTTAATCTCAGAGATCC 50740
QY 298 CCCAACGGAACCTTTGAGATTATATGCA-----ATCGAAGTGACAGGAAACATGCCAACT 352
DB 50741 TTAGCAGACAGCTGGATCGTCCATCCATCAGACTGAAGTGGCCGAGCAGATGCCAG 50800
QY 353 CAATCCCTTTAATGTACATGGA-TGGCCAAAGATGATTGGCAGCTCTCTTGCACAGTCG 411
DB 50801 CAATCCCTCTCGTGATACATGACGGGCGGAAGTCTTTGGCAGTCTCTAGTAGCCAG 50860
QY 412 ATGAGATGGA-GATGCTTGTCAATGAAGGGCCNCCTGTTGTCAATTCGAGTACAC 470
DB 50861 ATGGAGGTGGATGATGCTGTGCCCATGAAGGACCGGCTGTGGTCCCTTC-CGAGCTGC 50919
QY 471 AAAGAAAAAATGTCAATCCGAATCGAGGGAATATGCCCTTGGATTGCATGCTCTGCAG 530
DB 50920 GCAGGAAGAGAGCTTGGCCACAGCTGAGGGACACATGCCCTGGACTGATGTTCTGCAG 50979
QY 531 CCAGACCTTCACACATTCAGAAAGACTTTAATAAATGCTTTAATGCAACACCGGCCTAC 590
DB 50980 CCAGTCTTCTCTCAGCGGGAAGATCTCAATCAGCATGTGCTCCTTCGACACACCGCCAC 51039
QY 591 CCTCTGTGAACACAGAGTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAG 650
DB 51040 CCTCTGTGAGCGGCTGTGCTGCGTGTGGAGCCGAGTATTTAAGCCCTTCGACAAAG 51099
QY 651 TCAAGTGGGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAA---ATGAATTTAGCTG 707
DB 51100 TCAGTGCCCGGAGGACCAACCAAGAGAGAGCAGCAGAAAGAAATGGAGAGCTGAGCTG 51159
QY 708 TGAGGTATGCGGACAGCAATTTAGAGTCGCTTTTATGATGTTGAGATCCATCAGAACACA 767
DB 51160 TGATGATGCGACAGACATATCCCAAGTGCCTTTTATGATGTTGAGAGCCATCAAGAAACA 51219
QY 768 CAAGATTTCTTCACTTACGGGTGTAAATGTCGCGGAAGAAATTCAGAGGACCTTGGTT 827
DB 51220 TAAGGACTCTTTCAGTATGGGTGAGCATGTGCGGAGGAGATTCAGAGGACCGTGGTT 51279

```
QY 828 TCTTAAATAATCATGCGGACRCAATATGCGAAATCGGGGCGCAGAAAGCAAACTGCAGCA 887
Db 51280 CTTGAAGAACCACATCGGACACACAATGCGAAGTCTGGCACCAGCAAGCAAGCTGCAGCA 51339
QY 888 AGGCTTGGAGAGTAGTCCAGCAACGATCAACAGAGTCGTCCAGATGTTGGTCCAGGCACATGCCCTGGGAA 947
Db 51340 GGGCAGGAG---AGTCCAGTCAACATCAACGATGTTGGTCCAGGCACATGCCCTGGGAA 51396
QY 948 CATCTCTCTCTCTGCAAAATCTGCATGTTGTGTGTGCTTCTATTTCCAAATAAAGAAAG 1007
Db 51397 CATCTCCACTCCCTACAAGATCTGCATGTTGTGTGCTTCTCTTCCCAATAAAGCAGAG 51456
QY 1008 TCTAATTGACACCGCAAGGTGCACACCAAAAACCTGCTTTGCGTACAGCAGCGCCGCA 1067
Db 51457 CTTCAATTGACACAGCAAGGTTTCATCCAAAGAAACTGTTCTTAGTGCAGCAGCACAGC 51516
QY 1068 GACAGACTCTCCACAGGAGGATGCGCTCTCGAGGGAGGACTTCTCGAGTGTTCAA 1127
Db 51517 CCTGATGGCCACCAAGAGAACCCACATCCCGAGGGAAGAGCTCTCGAGTGTTCGAA 51576
QY 1128 CTTGAGACCAAAATCTCACCTGAAACGGGGGAAGAGCCTGTGCAGATGCATCCCTCAGCT 1187
Db 51577 CTTGAGACCCAGTCAACTGCAGAGAGTACAGTGAAGCCCATGACCTCTATACCTCAGCT 51636
QY 1188 CGATCGGTTACACACCTTCAGGCTTGGCAKCTGGCTACCAAGGAAGAGTGGCCATTG 1247
Db 51637 TGACCCATTCACCACTTACAGGATGCGAGTGGCTTGGCTACCAAGGCAAGAGTGGCCGTTGC 51696
QY 1248 CCAAGAA---GTGAGGAATTTGGGCAAGAGGAGCAGCAGCAGCAGATTCGAGTTC 1304
Db 51697 CCAAGAAAGGTGAAGAGTTCAGGCCCAAGAGGAGCAGCAGCAGATGATTCGGTTC 51756
QY 1305 CGAGA---AGGAGCTTGGAGAA----- 1323
Db 51757 AGAGAAGAGGAACTTGGAGAAATATGGTTGGGGTAAGTGGAGGTTCTGGAAGTC 51816
QY 1324 ----ACAAATAGAACCATTTGCGAGGCTCTCGCAAGAGAAAGAGAGTGCACACCTC 1379
Db 51817 CAAGACAAATAGAACGAGTGTGTTGAGTCTCTCCCAAGCAAAAGATAAGTCTAGACATGG 51876
QY 1380 CCACGCGAAGCGCCCTCGTGGAGCGCATCCCAAGTTACCAGTAGCAAGGAGAGCC 1439
Db 51877 TAACTAGTGAAGTGCCTTCAGGGAGCCTGACTCCAAAGTTGCCCGGAGCAAGAGAGCC 51936
QY 1440 CACTCACTCTCGAGTGGCGGAAAGCTTTGAGAACCTTACCAACAGCTGGTCTTGCACCTC 1499
Db 51937 CACACACTGTTCCGAGTGCAGTAAGCGTTTCAGGACATACCAAGCTGCTGCTGCACCTC 51996
QY 1500 CAGGGTCC 1507
Db 51997 GAGGGTGC 52004
```

```
RESULT 13
AL844576/c
LOCUS
DEFINITION Mouse DNA sequence from clone RP23-233M15 on chromosome 2, complete
sequence.
ACCESSION AL844576
VERSION AL844576.9 GI:27848013
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 114285)
Griffiths, C.
Direct Submission
Submitted (22-JAN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 22, 2003 this sequence version replaced gi:27820135.
Sequence from the Mouse Genome Sequencing Consortium whole genome
```

shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-233M15 is from the RPCI-23 Mouse BAC Library constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACE3.6.

FEATURES

1..114285
Location/Qualifiers
source
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-233M15"
/clone_lib="RPCI-23"

ORIGIN

Query Match 38.1%; Score 573.6; DB 10; Length 114285;
Best Local Similarity 65.6%; Pred. No. 1.3e-164;
Matches 1028; Conservative 3; Mismatches 448; Indels 88; Gaps 10;
QY 2 CAGGTGCTGGAGTTCAGTCTTCTGCTCAATTCGAAACATCTCAATGAGACAAAGAGC 61
Db 32266 CACATTGCGGGATCAACGTCTTGTCTCACTGAGCCACTCACAGTGGAGATGGACAGAG 32207
QY 62 ACTAATGCTTTGTGCTGATTTCATATTGAAATCGAGGCATTGGGAACCCCTGTATGCTTGT 121
Db 32206 CTGTTGCCCTCGTGTGACGCACATCTGACTCAGCGCTGGGCGAGCTTGGATGCTTGT 32147
QY 122 TTGTGGAAGAACCAAGTGCACCATCAGTCTCTTAAAGTTCCTTAAAGAGTTCGAAAGAGTTAGAG 181
Db 32146 TTGTGGAAGAGCCA---TCTCCATCACTGAGCTCTCCCAAGCTCC-----GG 32102
QY 182 ACTATACACTTCTTTTGGAACTTTTATAATAATATTGCTCTGTTTGGAAACCCAGGA 241
Db 32101 AATATACTCGTCCCTCTGACCTT-----AGAGGTGCTCTGATTTTGAACCCAGGG 32051
QY 242 CTGTTAGAGGGTGAG---TGACAGGTCTTACAGTGGCCCTTAATCCAACCTCCAGAAATTGC 298
Db 32050 CTCCTTAGAGTTCTCAGTGGTGATATCTCAGAGCGGCTTATTTCTAAGTCAGAGATCC 31991
QY 299 CCAACGGAACTTTGAGATTATATGCAATCGA-----AAGTGACAGGAACATGCCAACT 352
Db 31990 TCAGCAGATAGTGGATTACCCATCTCTGTGAGACCCCTGAGAGTGGTGGCAGCATGCCACC 31931
QY 353 CAATCCCTCTTAATGTACATGGA-TGGCCACAGAGTATGGCAGCTCTCTTTCAGCTCCG 411
Db 31930 CAGTCCCTCTCGTGTATATGACGCGGCGGAGTCTCTCAGCAGCTCTCTTAGGCTCCCGAG 31871

ORGANISM	Xenopus laevis	
REFERENCE	Eukaryota: Metazoa; Chordata: Craniata: Vertebrata: Euteleostomi: Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Xenopus.	
AUTHORS	1 (bases 1 to 3512) Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P.	
TITLE	Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative	
JOURNAL	Dev. Dyn. 225 (4), 384-391 (2002)	
PUBMED	12454917	
REFERENCE	2 (bases 1 to 3512) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Woxley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shervchenko,Y., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalka,U., Smalius,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.	
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
PUBMED	12477932	
REFERENCE	3 (bases 1 to 3512) Klein,S. and Strausberg,R.	
AUTHORS	Direct Submission	
TITLE	Submitted (01-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA	
JOURNAL	NIH-MGC Project	
REMARK	Contact: XGC help desk	
COMMENT	Email: cgabbs@mail.nih.gov Tissue Procurement: Dr. Igor Dawid cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu, Parvaneh Saeedi, JR Santos, Angeliqne Schnerch, Ursula Skalska, Duane Smaluis, Jeff Stott, Miranda Tsai, George Yang, Jacques Schein, Asim Siddiqui, Rob Holt, Marco Marra.	
FEATURES	source	
source	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 152 Row: m Column: 16 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein. Location/Qualifiers 1..3512 /organism="Xenopus laevis" /mol_type="mRNA" /db_xref="taxon:8355" /clone="MGC:80369 IMAGE:5074630" /tissue_type="Ovary, Xenopus" /clone_lib="NICHD_XGC_Ovl"	
gene	lab_host="DH10B" /note="Vector: pCMV-SPORT6" 1..3512 /gene="LOC398381" /db_xref="LocusID:398381" 210..3206 /gene="LOC398381" /codon_start=1 /product="LOC398381 protein" /protein_id="AAH72191.1" /db_xref="GI:47940272" /db_xref="LocusID:398381" /translation="MPVQSLSEFGDGVGSAIGSQMESPDSSLTKRTNTISHKNLH ETFLMQEGDLAFDCMFCDOYVKKHEDLQKVLVQHRPTLCEPEVLRVEAEVGLPLDK RKASQVLAKEVEEKEKEEFCEVCGQTFELSAVLGAQKHKHSQFTVCTTCGR RFKPFWLNKHKRTHTRGANKRPPLSDIELATINDVQDQKPKSATSPYKCMIC GFFPDKTTLMEHSMKHSKDSGLSDITNHSVPKNTRSSENCDDIAATVPKHFMS FINLNPTAHSKTVLDYKSGELDPFTYQWLATKGVALSGLRGKEPLLEPKAG ADKHGVKNADATLPHNHDKPGNCVKVPLFGKEEKLATKDPTLCSGCGKIPTYH ALVHSRVHNSRSDSSNLSGGEGALLVCAANVEVSHDQAVDPNPDPVARSVTLOE DSDEMDTQTDPEDEGEVRLKAGLPTSEKRCFGKTFRSNYLYNLHLRHTGKEP YRCFCDYAAAKQTLRLRHLRHHKFKPGDSNAHVKSIGKNSPTPGGTGDPMPVDNHV GNEARPKQTIINPTKDFVKAKDQCKRQPSSEGREBIIIPNRSTPTQVFSKEL PVPRSDQVPSPADWIMEENDSLCEKDHLLHNSLSIKEAQDNLSPMKEDVKPLNL CLKPSNSTPTNSALSVSTPCYSKSLYPEVLILHOKLTHKNSSEAVPKNIRKTA AOMAARBTGCLVLRGVDPVPLTYGKKCKTTPSPVKINPERKQLSFLPNKAAQS ESPPQGNKLVHTRHAPPLVNHSYMOPDRGIIHLLERMQSQSNVTARDITYGRN SSSVFMSGSEYFFSRMLNLSQESGEPYSKRKKNTPSADSTVYPTKTDVILNTPYSK LASFLLOEKLNSGAPLLPHSASSHSAVESWRPPIIPAPPTAGTYRALDPPSFSREP SSITEAKRAPINRHVSXRVFGPNEKA"	
ORIGIN	Query Match 14.5%; Score 217.8; DB 5; Length 3512; Best Local Similarity 62.3%; Pred. No. 4.6e-55; Matches 367; Conservative 0; Mismatches 207; Indels 15; Gaps 1; QY 476 AAAAAATGTCGAATCGAATCGAGGGGAATATGATGCCCTTCGATTGCAATGTTCTGCACCCAGA 535 Db 343 AAACCTTTTAAATGACGACGAGGAGACCTTGGCATTGACTGCATGTTTTCGACACCA 402 QY 536 CTTTACACATTCAGAGACCTTAAATAACATGCTTAAATGCAACACCGCGCTACCCCTCT 595 Db 403 CCTACAAGCATCAGAGGATCTAGGAAAACATGATTGTTGTCACCATAGGCGCAACCCCTCT 462 QY 596 GTGACACGAGCTTCTTCGGGTGAGCAGAGTAGTCTCAGTCCGCTTGTATATAAATGTCAC- 654 Db 463 GTGAACCAAGAGTCTCTCGGTGTTAGGCGGAGTAGCTTGGTCTCTCTGATATAGCGCCGAA 522 QY 655 -----GTGCGAACAGAACCTCCCAAGGAAAAGAAATTTGCAAGGAAAAATGAAT 700 Db 523 AAGCTTCTGTGCTCTGTTGTAAGAGAGGTAAAGAGAGATGAAGAGAGGAGGAGT 582 QY 701 TTAGCTGTGAGGTATGTGGGAGACATTTAGAGTGCCTTTTGTAGTGTGAGATGCCACATGA 760 Db 583 TTGACTGTGAGGTTGTGCTCAGACTTTTGAAGCTTCTGCTGCTCTTGGGTGCCACCA 642 QY 761 GAACACACAAGATCTTTTCACTTACGGGTGTACATGTGCGGAGAGAGATTCAGGAGC 820 Db 643 AGAAGCACAAGATTCCTTCACTTACTCTGCTACTCTGTTGTTGTTGTTGTTGTTGTTGTT 702 QY 821 CTTGTTTCTTTAAAAATCACATCGGACRCATATATGGCAAAATCGGGGGCCAGAACCAAC 880 Db 703 CTGTTTCTTTAAAAATCATTAAGAGAACCCATACAAACAGATCAGGGGCCAAAAACAAAC 762 QY 881 TGCAGCAAGGCTTGAGAGTAGTTCAGCAACGATCAACGAGGTCTCTCAGGTGTCAGCGG 940 Db 763 GCCTCTTTCTTGATGAAGAGATACCTGGCGGACAAATAAATCATGTTGTTCAAGATCAGAAC 822 QY 941 CCGAGGACATCTCTCTCTTTCGAAAATCTCATGGTTTGTGGCTTCTCTATTTTCCAAATA 1000 Db 823 CCAAAAGCGCAACATCTCTCTTATAAATTTGTGATGATATGTTGGGTCTCTCTTCCCTGATA 882 QY 1001 AAGAAAGTCTAATTGAGCACCGCAGGTGTCACACCAAAAAAACTGCTTT 1049	

Db 883 AATGCACCCCTCATGGAACATAGTAGATGCATTCCAAGGATTCTGGTTT 931

Search completed: June 29, 2005, 01:05:49
Job time : 4340.72 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:28:52 ; Search time 527.33 Seconds
(without alignments)
16917.381 Million cell updates/sec

Title: US-08-731-499-3

Perfect score: 1507

Sequence: 1 GCAGTTCTGGATTGACT.....GGTCTTGCACTCCAGGGTCC 1507

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1504.8	99.9	1507	2	Aav04698 Homo sapi
2	1387.2	92.1	36022	11	Acn44986 Human gen
3	1365	90.6	10282	2	Aav09023 Homo sapi
4	1337.2	88.7	5632	8	Acc50334 Breast ca
5	1337.2	88.7	5632	10	Add14635 Human src
6	1337.2	88.7	5632	13	Adr66761 Human pro
7	1337.2	88.7	5632	13	Adr65858 Human pro
8	1337.2	88.7	5632	12	Adp07267 Human znf
9	1132.6	75.2	5252	11	Acn44987 Human mrn
10	1087.2	72.1	3183	2	Aav09024 Homo sapi
11	573.6	38.1	26345	11	Acn44984 Mouse gen
12	537	35.6	3016	11	Acn44985 Mouse mrn
13	523.2	34.7	566	4	Aak81095 Human imm
14	324.2	21.5	14906	4	Aak81093 Human imm
15	215.6	14.3	530	12	ACH76516 Human gen
16	197.2	13.1	298	3	Aac07753 Human src
17	156.4	10.4	469	2	Aav88042 EST clone
18	104.8	7.0	5994	13	Acn40355 Tumour-as
19	103.2	6.8	6033	13	Adr25650 Breast ca
20	103.2	6.8	8156	6	Abt07614 Human bre

c	21	97	6.4	251	12	ACH90216	ACH90216 Human gen
	22	70	4.6	70	13	ADR00072	ADR00072 ZNF217 PC
	23	70	4.6	2706	10	ACH07836	ACH07836 Novel cod
c	24	68.8	4.6	1014	12	ACH87715	ACH87715 Human gen
	25	68.8	4.6	1229	12	ACH86998	ACH86998 Human gen
c	26	68.8	4.6	2706	11	ADL71050	ADL71050 Gene enco
	27	68.8	4.6	2765	4	AAI58185	AAI58185 Human pol
	28	68.8	4.6	2765	5	ADQ98391	ADQ98391 DNA enco
	29	68.8	4.6	2765	9	ADB48151	ADB48151 Novel hum
	30	68.8	4.6	3039	13	ADR25783	ADR25783 Breast ca
	31	68.8	4.6	3040	13	ADQ87590	ADQ87590 Human tum
	32	68.8	4.6	3617	4	ABK42507	ABK42507 Genomic s
	33	68.8	4.6	3617	4	ABK42509	ABK42509 Genomic s
	34	68.8	4.6	3617	9	ADB60663	ADB60663 Connectiv
	35	68.8	4.6	3617	9	ADB60665	ADB60665 Connectiv
	36	63.6	4.2	2909	11	ADL71048	ADL71048 Gene enco
	37	59.2	3.9	508	12	ACH73808	ACH73808 Human gen
	38	59.2	3.9	772	4	AAS31069	AAS31069 Human dia
	39	59.2	3.9	2375	11	ADM02211	ADM02211 Human cdn
	40	59.2	3.9	10926	4	AAS25370	AAS25370 Human imm
	41	58.4	3.9	1995	4	AAS25920	AAS25920 Human cdn
	42	58.4	3.9	1995	8	ABX73261	ABX73261 Human nov
	43	53.6	3.6	1584	4	AAS26371	AAS26371 Human cdn
	44	53.6	3.6	1584	8	ABX73712	ABX73712 Human nov
c	45	53.6	3.6	2014	4	AAS25921	AAS25921 Human cdn

ALIGNMENTS

RESULT 1

AAV04698

ID AAV04698 standard; cdna; 1507 BP.

XX AAV04698;

AC AAV04698;

DT 21-JUL-1998 (first entry)

XX Homo sapiens 20q13 amplicon cc49 transcript.

XX 20q13 amplicon; chromosome 20; tumour; detection; C2H2 zinc finger;
XX chromosomal abnormalities; probe; gene therapy; antisense inhibition;
XX treatment; age-related macular degeneration; retinitis pigmentata;
XX Leber's congenital anaurosis; ds.

OS Homo sapiens.

XX WO9802539-A1.

PN WO9802539-A1.

XX 22-JAN-1998.

XX 15-JUL-1997; 97WO-US012343.

XX 15-JUL-1996; 96US-00680395.

PR 16-OCT-1996; 96US-00731499.

PR 17-JAN-1997; 97US-00785532.

XX (REGC) UNIV CALIFORNIA.

XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;

WPI; 1998-110587/10.

XX New sequences from the 20q13 amplicon - used for detecting chromosomal

XX abnormalities, particularly tumours, and for developing products for

XX treating diseases.

XX Claim 1; Page 61-62; 91pp; English.

XX The sequence is that of a cdna sequence cc49, which was isolated from the

XX 20q13 amplicon. It shows homology to C2H2 zinc finger genes. It can be

XX used as a probe for the detection of chromosomal abnormalities at 20q13.

XX It and other sequences isolated from the 20q13 amplicon are consistently

CC amplified in primary tumours. These sequences are useful as probes or as
CC probe targets for monitoring the relative copy number of corresponding
CC sequences from a biological sample such as tumour cells. The sequences
CC can also be used in therapeutic applications for modulating the
CC expression of the endogenous gene or the activity of the gene product.
CC Examples of therapeutic approaches include antisense inhibition of gene
CC expression, gene therapy, and monoclonal antibodies that specifically
CC bind the gene products. The products can also be used in the treatment of
CC other diseases, e.g. age-related macular degeneration, Leber's congenital
CC amaurosis and retinitis pigmentation
XX

SQ Sequence 1507 BP; 445 A; 354 C; 364 G; 340 T; 0 U; 4 Other;

Query Match 99.9%; Score 1504.8; DB 2; Length 1507;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCAGGTTCGCGAATGACTTCTTGCTCAATTAAGAACACTCAATCGAGCAAAAGAG 60
Db 1 GCAGGTTCGCGAATGACTTCTTGCTCAATTAAGAACACTCAATCGAGCAAAAGAG 60
Qy 61 CACTAATGCTTGCTGATTCATATTTGAATCGAGGCAATGGGAACCCGTATGCCCTTG 120
Db 61 CACTAATGCTTGCTGATTCATATTTGAATCGAGGCAATGGGAACCCGTATGCCCTTG 120
Qy 121 TTTGTGGAAAGAACAGTGACACCATCACTGAGCTTCTTAAAGTTTCGAAGATTAGAG 180
Db 121 TTTGTGGAAAGAACAGTGACACCATCACTGAGCTTCTTAAAGTTTCGAAGATTAGAG 180
Qy 181 GACTATACACTTTCTTTTGAACCTTTTATAATAAATATTTGCTCTGGTTTGGAAACCCAGG 240
Db 181 GACTATACACTTTCTTTTGAACCTTTTATAATAAATATTTGCTCTGGTTTGGAAACCCAGG 240
Qy 241 ACTGTTAGAGGTGAGTGACAGGCTTACAGTGGCTTTAATCCAACTCCAGAAATGGCC 300
Db 241 ACTGTTAGAGGTGAGTGACAGGCTTACAGTGGCTTTAATCCAACTCCAGAAATGGCC 300
Qy 301 AACGGAACTTTGAGATTATATGCAATCGAAGTGCAGGAAACATGCCAATCAATCCCT 360
Db 301 AACGGAACTTTGAGATTATATGCAATCGAAGTGCAGGAAACATGCCAATCAATCCCT 360
Qy 361 CTTAATGTACATGATGCGCAAGAGTGTGTCAGCTCTCTTGCCAGTCCGATGAGATG 420
Db 361 CTTAATGTACATGATGCGCAAGAGTGTGTCAGCTCTCTTGCCAGTCCGATGAGATG 420
Qy 421 GAGATGCCCTTGTCAATGAAAGGCCCNCTGTGTCAATTCGAGCTACACAAAGAAAAA 480
Db 421 GAGATGCCCTTGTCAATGAAAGGCCCNCTGTGTGTCAATTCGAGCTACACAAAGAAAAA 480
Qy 481 ATGTCAAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCGAGCCAGACCTTC 540
Db 481 ATGTCAAATCCGAATCGAGGGGAATATATGCCCTTGGATTGCATGTTCTGCGAGCCAGACCTTC 540
Qy 541 ACACATTGAGAGACCTTTAATAAATGCTTTAATGCAACACCCGCCCTACCTCTGTGAA 600
Db 541 ACACATTGAGAGACCTTTAATAAATGCTTTAATGCAACACCCGCCCTACCTCTGTGAA 600
Qy 601 CCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGGGA 660
Db 601 CCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGGGA 660
Qy 661 ACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGTGGG 720
Db 661 ACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGTGGG 720
Qy 721 CAGACATTTAGAGTCCGCTTTTGTGATGTTGAGATCCACATGAGAACACAAAGATTCTTTC 780
Db 721 CAGACATTTAGAGTCCGCTTTTGTGATGTTGAGATCCACATGAGAACACAAAGATTCTTTC 780
Qy 781 ACTTACGGGTGTAAATGTGCGGAAGAAATTCAGGAGCCCTTGGTTTCTTAAAAATCAC 840
Db 781 ACTTACGGGTGTAAATGTGCGGAAGAAATTCAGGAGCCCTTGGTTTCTTAAAAATCAC 840

Qy 841 ATCGGACRCATATATGGCAAAATCGGGGCCAGAGCAAAATGCGAGCAAGGCTTGAGAGT 900
Db 841 ATCGGACRCATATATGGCAAAATCGGGGCCAGAGCAAAATGCGAGCAAGGCTTGAGAGT 900
Qy 901 AGTCCAGCAACGATCAACGAGGTGCTCCAGGTGACGCGGCGGAGAGCATCTCTCTCT 960
Db 901 AGTCCAGCAACGATCAACGAGGTGCTCCAGGTGACGCGGCGGAGAGCATCTCTCTCT 960
Qy 961 TGCAAAATCTGATGTTTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCAC 1020
Db 961 TGCAAAATCTGATGTTTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCAC 1020
Qy 1021 CGCAAGGTGCACACCAAAAAAATGCTTTTCGTACACGAGCGCGCAGACACTCTCCA 1080
Db 1021 CGCAAGGTGCACACCAAAAAAATGCTTTTCGTACACGAGCGCGCAGACACTCTCCA 1080
Qy 1081 CAAGGAGGAATGCCGTCTCCAGGGAGGACTTCTGCAAGTTGTTCAACTTGAGACCAAAA 1140
Db 1081 CAAGGAGGAATGCCGTCTCCAGGGAGGACTTCTGCAAGTTGTTCAACTTGAGACCAAAA 1140
Qy 1141 TCTCAACCTGAAACGGGGAAGAGCCTGTGATGTCATCCCTCAGCTCGATCCGTTCA 1200
Db 1141 TCTCAACCTGAAACGGGGAAGAGCCTGTGATGTCATCCCTCAGCTCGATCCGTTCA 1200
Qy 1201 ACCTTCCAGGCTTGCGCAKCTGGCTACCAAGGAAGTTGCGCATTTGCCAAGAGTGAAG 1260
Db 1201 ACCTTCCAGGCTTGCGCAKCTGGCTACCAAGGAAGTTGCGCATTTGCCAAGAGTGAAG 1260
Qy 1261 GAATTTGGGCGCAAGAGGAGCACCGCAACGACGATTCGAGTTCGGAAGAGGCTTGA 1320
Db 1261 GAATTTGGGCGCAAGAGGAGCACCGCAACGACGATTCGAGTTCGGAAGAGGCTTGA 1320
Qy 1321 GAAACAAATAGAACCAATTTGCGAGGCTCTCGCAAGGAAGAGAGTGCACAACTCC 1380
Db 1321 GAAACAAATAGAACCAATTTGCGAGGCTCTCGCAAGGAAGAGAGTGCACAACTCC 1380
Qy 1381 CACGCGAAGCGCCCTCCGTTGAGCGCGATCCCAAGTTACCCAGTAGCAAGGAGAGGCC 1440
Db 1381 CACGCGAAGCGCCCTCCGTTGAGCGCGATCCCAAGTTACCCAGTAGCAAGGAGAGGCC 1440
Qy 1441 ACTCACTGCTCCGAGTGGGCAAGAGCTTTTCAGAACTTACCACTCCAGCTGGTCTTGAC 1500
Db 1441 ACTCACTGCTCCGAGTGGGCAAGAGCTTTTCAGAACTTACCACTCCAGCTGGTCTTGAC 1500
Qy 1501 AGGGTCC 1507
Db 1501 AGGGTCC 1507
RESULT 2
ACN44986
ID ACN44986 standard; DNA; 36022 BP.
XX
AC ACN44986;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG37127.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SBP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX

KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentata;
KW Leber's congenital amaurosis; zinc finger amplified in breast cancer; ds.
XX
XX Homo sapiens.
XX
XX WO9802539-A1.
XX
XX 22-JAN-1998.
XX
XX 15-JUL-1997; 97WO-US012343.
XX
XX 15-JUL-1996; 96US-00680395.
XX 16-OCT-1996; 96US-00731499.
XX 17-JAN-1997; 97US-00785532.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX
XX WPI; 1998-110587/10.
XX P-PSDB; AAW23975.
XX
XX New sequences from the 20q13 amplicon - used for detecting chromosomal
XX abnormalities, particularly tumours, and for developing products for
XX treating diseases.
XX
XX Claim 1; Page 64-67; 91pp; English.
XX
XX The sequence is that of the genomic sequence of ZABC-1 (zinc finger
XX amplified in breast cancer). It maps to the core of the 20q13.2 amplicon
XX and is overexpressed in primary tumours and breast cancer cell lines
XX having 20q13.2 amplification. The exact coding region for the genomic
XX sequence is not given. The sequence can be used as a probe for the
XX detection of chromosomal abnormalities at 20q13. It and other sequences
XX isolated from the 20q13 amplicon are consistently amplified in primary
XX tumours. These sequences are useful as probes or as probe targets for
XX monitoring the relative copy number of corresponding sequences from a
XX biological sample such as tumour cells. The sequences can also be used in
XX therapeutic applications for modulating the expression of the endogenous
XX gene or the activity of the gene product. Examples of therapeutic
XX approaches include antisense inhibition of gene expression, gene therapy,
XX and monoclonal antibodies that specifically bind the gene products. The
XX products can also be used in the treatment of other diseases, e.g. age-
XX related macular degeneration, Leber's congenital amaurosis and retinitis
XX pigmentata
XX
XX Sequence 10282 BP; 2820 A; 2222 C; 2191 G; 3047 T; 0 U; 2 Other;
XX
XX Query Match 90.6%; Score 1365; DB 2; Length 10282;
XX Best Local Similarity 96.4%; Pred. No. 0;
XX Matches 1455; Conservative 4; Mismatches 45; Indels 6; Gaps 6;
XX
XX 2 CAGGTTGCTGGATGACTCTTGTCTCAATTGAACACTCATTCAATGAGACAAAGAGC 61
XX
XX 2423 CAGGTTGCTCGAATGACTCTTGTCTCAATTGAACACTCATTCAATGAGACAAAGAGA 2482
XX
XX 62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCCTGTATGCTTGT 121
XX
XX 2483 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCCTGTATGCTTGT 2542
XX
XX 122 TTGTGGAAGAACCACTGACACCATCAGCTTCTTAAAGATTCGAAGAAGTTAGAGG 181
XX
XX 2543 TTGTGGAAGAACCACTGACACCATCAGCTTCTTAAAGATTCGAAGAAGTTAGAGG 2602
XX
XX 182 ACTATACACTTCTTTTGAACCTTTTATAATAATATTTGCTCTGCTTTTGAACCCAGCA 241
XX
XX 2603 ACTATACACTTCTTTTGAACCTTTTATAATAATATTTGCTCTGCTTTTGAACCCAGG 2662
XX
XX 242 CTGTTTGA-GGGTGAGTGACAGGCTTAC-AGTGGCCTTAATCCAACTCCAGAAATGCC 299
XX
XX 2663 CTGTTTGAAGGGGTGAGTGACAAAGTCTTACAAAGTGGCCTTATTTCCAACTCCAGAAATGCC 2722

QY 300 CAACGGAATTTGAGATTATATGCAATCGAAAGTCACAGGAAACATGCCAACTCAATCCC 359
DB 2723 CAACGGAATTTGAGATTATATGCAATCGAAAGTCACAGGAAACATGCCAACTCAATCCC 2782
QY 360 TCTTAATGTACATGGAT-GGCAAGAGTGATTGGCAGCTCTCTTCCAGTCCGATGGAGA 418
DB 2783 TCTTAATGTACATGGATGGGCCAGAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGA 2842
QY 419 TGGG-GATGCCCTTGTCAATGAAAGGCCCNCTGTTGTCAATTCGAGCTACACAAAGAAA 477
DB 2843 TGGAGGATGCCCTTGTCAATGAAAGGCCCGCTGTTGTTCCATTCGAGCTACACAAAGAAA 2902
QY 478 AAAATGCTCAATCCGNAATCGAGGGGNAATATGCCCTTGGATTGCAATGTTCTGAGCCAGACC 537
DB 2903 AAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCAATGTTCTGAGCCAGACC 2961
QY 538 TTCACACATTCAGAAAGACCTTAAATAAACATGCTCTTAATGCAACACACGGCCTTACCCTCTGT 597
DB 2962 TTCACACATTCAGAAAGACCTTAAATAAACATGCTCTTAATGCAACACACGGCCTTACCCTCTGT 3021
QY 598 GAACGAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTTGATAAAAGTCAAGTG 657
DB 3022 GAACGAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTTGATAAAAGTCAAGTG 3081
QY 658 CGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAAATTTAGCTGTGAGTATGT 717
DB 3082 CGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAAATTTAGCTGTGAGTATGT 3141
QY 718 GGGCAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTCT 777
DB 3142 GGGCAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTCT 3201
QY 778 TTCACTTACGGGTGTAAACATGTGCGGAAGAGATTCGAAGGAGCCTTTGTTCTTAAAAAT 837
DB 3202 TTCACTTACGGGTGTAAACATGTGCGGAAGAGATTCGAAGGAGCCTTTGTTCTTAAAAAT 3261
QY 838 CACATCGGACRCATATGCGCAAAATCGGGGCCAGAACCAACATCGACAGAGCTTGGAG 897
DB 3262 CACATCGGACRCATATGCGCAAAATCGGGGCCAGAACCAACATCGACAGAGCTTGGAG 3321
QY 898 AGTAGTCCAGCAACATCAACGAGGTCTCTCAGGTGACGCGGCCGAGAGCATCTCTCT 957
DB 3322 AGTAGTCCAGCAACATCAACGAGGTCTCTCAGGTGACGCGGCCGAGAGCATCTCTCT 3381
QY 958 CTTTGCAAAATCTGCAATGGTTTGTGCTTCTATTTTCAAAATAAAGAAAGTCTAATTTAG 1017
DB 3382 CTTTGCAAAATCTGCAATGGTTTGTGCTTCTATTTTCAAAATAAAGAAAGTCTAATTTAG 3441
QY 1018 CACCGAAGGTGCACACCAAAAACATGCTTTTCGGTACCAGCAGCGCCAGACACACTCT 1077
DB 3442 CACCGAAGGTGCACACCAAAAACATGCTTTTCGGTACCAGCAGCGCCAGACACACTCT 3501
QY 1078 CCACAAAGGAGGAATGCCCTCTCGAGGGAGGACTTCTTGCAGTGTTCCTCACTTGAGACCA 1137
DB 3502 CCACAAAGGAGGAATGCCCTCTCGAGGGAGGACTTCTTGCAGTGTTCCTCACTTGAGACCA 3560
QY 1138 AAATCTCACTCCCTGAAAACGGGGAAAGAGCCTGTGAGATGCAATCCCTCAGTGCATCCGTT 1197
DB 3561 AAATCTCACTCCCTGAAAACGGGGAAAGAGCCTGTGAGATGCAATCCCTCAGTGCATCCGTT 3620
QY 1198 ACCACCTTCCAGGCTTGGCACTGGCTACCAAGAAAGTTCGCTATTTGCCAAGAGTG 1257
DB 3621 ACCACCTTCCAGGCTTGGCACTGGCTACCAAGAAAGTTCGCTATTTGCCAAGAGTG 3680
QY 1258 AAGGAATTTGGGGCAAGAAAGGGAGCACCGCAACCGAGTTCGAGTTCGAGAGAGAGCTT 1317
DB 3681 AAGGAATTTGGGGCAAGAAAGGGAGCACCGCAACCGAGTTCGAGTTCGAGAGAGAGCTT 3740
QY 1318 GGAGAAAACAAATAGAACCATTTGTGAGGCTCTTCGAAAGAGAAAGAGAGTGCACACAC 1377
DB 3741 GGAGAAAACAAATAGAGGAGTGTGTGAGGCTCTTCGAAAGAGAAAGAGAGTGCACACAC 3800
QY 1378 TCCACGGCGAAGGCCCTCTCGTGGAGCGGGATCCCAAGTTTACCAGTAGCAGAGAGAG 1437

Db 3801 TCCACGGCGAAGCGCCCTCGGTGACGCGGATCCCAAGTTTACCAGTAGCAGGAGG 3860
 Qy 1438 CCACCTCACTGCTCCGAGTGCGGCAAGCTTTCAAGAACCTACACAGCTGCTTTGCAC 1497
 Db 3861 CCACCTCACTGCTCCGAGTGCGGCAAGCTTTCAAGAACCTACACAGCTGCTTTGCAC 3920
 Qy 1498 TCCAGGGTCC 1507
 Db 3921 TCCAGGGTCC 3930
 RESULT 4
 ID ACC50334
 AC ACC50334 standard; cDNA; 5632 BP.
 AC ACC50334;
 XX 12-JUN-2003 (first entry)
 DE Breast cancer associated cDNA sequence SEQ ID NO:505.
 DE Human; breast cancer; cytostatic; gene therapy; gene; ss.
 KW Homo sapiens.
 OS Homo sapiens.
 XX WO2003004989-A2.
 PN 16-JAN-2003.
 PD 21-JUN-2002; 2002WO-US019669.
 XX 21-JUN-2001; 2001US-0299887P.
 PR 27-JUN-2001; 2001US-0301572P.
 PR 18-JUL-2001; 2001US-0306501P.
 PR 25-SEP-2001; 2001US-0325002P.
 PR 05-MAR-2002; 2002US-0362585P.
 PR 14-MAY-2002; 2002US-0380391P.
 XX (MILL-) MILLENIUM PHARM INC.
 XX Lillie J, Gannavarapu M, Glatt K, Hoerssh S, Kamathkar S;
 PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
 PI Bast RC, Hortobagyi GN, Pusztai L, Mexic F, Sahin A, Mills GB;
 XX WPI: 2003-210381/20.
 DR P-PSDB; ABR47632.
 DR Breast cancer diagnosis or treatment by comparing the level of expression
 PT of a marker in a patient sample with that in the control non-breast
 PT cancer sample.
 XX Claim 1; SEQ ID NO 505; 128pp; English.
 XX The present invention describes a method for assessing whether a patient
 CC is afflicted with breast cancer. The method comprises comparing the level
 CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
 CC ABR47386 to ABR47632) in a patient sample and the normal level of
 CC expression of the marker in a control non-breast cancer sample, where a
 CC significant increase in the level of expression of the marker in the
 CC patient sample and the normal level is an indication that the patient is
 CC afflicted with breast cancer. The breast cancer associated sequences from
 CC the present invention have cytostatic activities and can be used in gene
 CC therapy. The method is useful for diagnosing and treating breast cancer.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;
 SQ

Query Match 88.7%; Score 1337.2; DB 8; Length 5632;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

Qy 52 GACAAAGAGCACTAATGCTTTGCTGATTCATATTTGAATCAGGCAATTTGGAAACCTTG 111
 Db 1 GACAAAGAGCACTAATGCTTTGCTGCTGATTCATATTTGAATCAGGCAATTTGGAAACCTTG 60
 Qy 112 TATGCTTTGTTGTGGAAGAACAGAGTGACACCATCAGTGAGCTTCTTAAAGTTTGAAG 171
 Db 61 TATGCTTTGTTGTGGAAGAACAGAGTGACACCATCAGTGAGCTTCTTAAAGTTTGAAG 120
 Qy 172 AAGTTAGAGGACTATACACTTTCTTTTGAATTTTATATAATAATATTTGCTCTGG-TTTT 230
 Db 121 AAGTTAGAGGACTATACACTTTCTTTTGAATTTTATATAATAATATTTGCTCTGGTTT 180
 Qy 231 GGAACCCAGGACTGTGTAGA-GGGTGAGTGACAGGCTTTAC-AGTGGCCCTTAATCCAACCTC 288
 Db 181 GGAACCCAGGCTGTTAGAGGGGTGAGTGACAGGCTTTACAAGTGCCCTTATTTCCAACCTC 240
 Qy 289 CAGAAATTTGCCCAACGGAACCTTTGAGATTATATGCAATCGAAGTGACAGGAAACATGCG 348
 Db 241 CAGAAATTTGCCCAACGGAACCTTTGAGATTATATGCAATCGAAGTGACAGGAAACATGCG 300
 Qy 349 AACTCAATCCCTCTTAATGTATCATGGAT-GGCCAAGAGTGATTTGGCAGCTCTCTTGCCAG 407
 Db 301 AACTCAATCCCTCTTAATGTATCATGGATGGGCCAGAGTGATTTGGCAGCTCTCTTGCCAG 360
 Qy 408 TCCGATGGAGATGGA-GATGCTTCTCAATGAAAGGCCCCNCTGTTGTCAATTTCCGAGCT 466
 Db 361 TCCGATGGAGATGGAGGATGCCCTTCTCAATGAAAGGCCCGCTGTTGTTCCATTTCCGAGC 420
 Qy 467 ACACAAAGAAAAAATGTCATCCGAATCGAGGGGAATATGCCCCCTTGGATTTGATGCTTCT 526
 Db 421 TACACAAAGAAAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTTGATGCTTCT 479
 Qy 527 GCAGCCAGACCTTTCACATTCAGAAAGACCTTTAAATAACATGCTCTTAATGCAACACCGGC 586
 Db 480 GCAGCCAGACCTTTCACATTCAGAAAGACCTTTAAATAACATGCTCTTAATGCAACACCGGC 539
 Qy 587 CTACCTCTGTGAAACCAGCAGTTCTTCGGGTTTGAAGCAGAGTATCTCAGTCCGCTTGATA 646
 Db 540 CTACCTCTGTGAAACCAGCAGTTCTTCGGGTTTGAAGCAGAGTATCTCAGTCCGCTTGATA 599
 Qy 647 AAGTCAAGTGGGAACAGACCTCCCAAGGAAGAAAGATTTGCAAGGAATGATTTAGCT 706
 Db 600 AAGTCAAGTGGGAACAGACCTCCCAAGGAAGAAAGATTTGCAAGGAATGATTTAGCT 659
 Qy 707 GTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACAC 766
 Db 660 GTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACAC 719
 Qy 767 ACACAAAGATTTCTTTCACTTACGSGGTGTAACATGTGCGGAAGAAAGATTTCAAGGAGCTTGGT 826
 Db 720 ACACAAAGATTTCTTTCACTTACGSGGTGTAACATGTGCGGAAGAAAGATTTCAAGGAGCTTGGT 779
 Qy 827 TTTCTTAAAAATCAGATGGGACRCAATATGGAATCGGGGCCAGAGCAACATGTCAGC 886
 Db 780 TTTCTTAAAAATCAGATGGGACRCAATATGGAATCGGGGCCAGAGCAACATGTCAGC 839
 Qy 887 AAGGCTTCGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGACGCGGCCGAGA 946
 Db 840 AAGGCTTCGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGACGCGGCCGAGA 899
 Qy 947 GCATCTCTCTCTTTTCAAAAATCTGCATGGTTTGTGGCTTCTTATTTCCAAATAAAGAAA 1006
 Db 900 GCATCTCTCTCTTTTCAAAAATCTGCATGGTTTGTGGCTTCTTATTTCCAAATAAAGAAA 959
 Qy 1007 GTCTAATTTAGCACCGCAAGGTGCAACCAAAAAAATGCTTTTCGGTACACAGGCGGC 1066
 Db 960 GTCTAATTTAGCACCGCAAGGTGCAACCAAAAAAATGCTTTTCGGTACACAGGCGGC 1019
 Qy 1067 AGACAGACTCTCCCAAGGAGGAATGCCGCTCTCCAGGGAGGACTTCTTCGAGTTGTTCA 1126
 Db 1020 AGACAGACTCTCCCAAGGAGGAATGCCGCTCTCCAGGGAGGACTTCTTCGAGTTGTTCA 1079

Qy	1127	ACTTGAGACCAAAATCTCA	CCCTGAAACGGGGAAAGACCTGTGCAGATGCATCCCTCAGC	1186
Db	1080			
		ACTTGAGACCAAAATCTCA	CCCTGAAACGGGGAAAGACCTGTGCAGATGCATCCCTCAGC	1139
Qy	1187	TGCATCCGTTCCACCACCTTC	CCAGCGCTTGGCAKCTGGCTACCAAGGAAGAGTTGCCATTT	1246
Db	1140			
		TGCATCCGTTCCACCACCTTC	CAGCGCTTGGCAGCTGGCTACCAAGGAAGAGTTGCCATTT	1199
Qy	1247	GCCAAGAAGTGAAGGAATTT	GGGGCAAGAAAGGAGCACCGACAACGACGATTTCCAGATTCCG	1306
Db	1200			
		GCCAAGAAGTGAAGGAATTT	GGGGCAAGAAAGGAGCACCGACAACGACGATTTCCAGATTCCG	1259
Qy	1307	AGAAGGAGCTTGGAGAAACAAA	TAAGAACCAATTTGTGCAGGCGTCTCCGACAGAAAGAGA	1366
Db	1260			
		AGAAGGAGCTTGGAGAAACAAA	TAAGGGCGAGTTGTGCAGGCGTCTCCGACAGAAAGAGA	1319
Qy	1367	AGTGCAAACTCCACCGGACGAGCGCCCT	CCGTGGACGCGATCCCAAGTTACCCAGTA	1426
Db	1320			
		AGTGCAAACTCCACCGGACGAGCGCCCT	CCGTGGACGCGATCCCAAGTTACCCAGTA	1379
Qy	1427	GCAAGGAGAAGCCCACTCACT	GTCTCCAGTGCGGCAAAAGCTTTTCAGAACCTACCAACGAGC	1486
Db	1380			
		GCAAGGAGAAGCCCACTCACT	GTCTCCAGTGCGGCAAAAGCTTTTCAGAACCTACCAACGAGC	1439
Qy	1487	TGGTCTTGCATC	CAGGGTCC	1507
Db	1440			
		TGGTCTTGCATC	CAGGGTCC	1460

RESULT 5

RESULTS 3
ADD14635
ID ADD14635 standard; cDNA; 5632 BP.

01-JAN-2004 (first entry)

Human src biomarker polynucleotide SEQ ID NO:29.

predictor set; protein tyrosine kinase activity modulator;
 KW
 protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
 KW
 gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
 KW

OS Homo sapiens.

XX PN WO2003062395-A2.

XX
PD
31-JUL-2003.

17-JAN-2003: 2003WO-US001981.

XX
PR 18-JAN-2002: 2002US-0350061P.XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.XX
PI Huang F, Fairchild CR, Lee FY, Shaw P;XX
DR WPI: 2003-636735/60.

DR WP1; 2003-636735/
DR P-PSDB; ADD14040;

PT New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.

XX
PS
Claim 2; SEQ ID NO 29; 139pp; English.

xx The present invention describes a predictor set comprising a plurality of
cc polynucleotides or polypeptides whose expression pattern is predictive of
cc the response of cells to treatment with a compound that modulates protein
cc tyrosine kinase activity or members of the protein tyrosine kinase
cc pathway. Also described: (1) predicting whether a compound is capable of
cc modulating the activity of cells, comprising obtaining a sample of cells,
cc determining whether the cells express a plurality of markers, and
cc

correlating the expression of the markers to the compound's ability to modulate the activity of the cells; (2) a plurality of cell lines for identifying polynucleotides and polypeptides whose expression levels correlate with compound sensitivity or resistance of cells associated with a disease state; and (3) identifying polynucleotides and polypeptides that predict compound sensitivity or resistance of cells associated with a disease state, comprising subjecting the plurality of cell lines to one or more compounds, analysing the expression pattern of a microarray of polynucleotides or polypeptides, and selecting polynucleotides or polypeptides that predict the sensitivity or resistance of cells associated with a disease state by using the expression pattern of the microarray. The polynucleotides and polypeptides have cytostatic activities, and can be used in gene therapy. The polynucleotides and polypeptides are useful in predicting the activity of compounds that interact with protein tyrosine kinases and/or protein tyrosine kinase pathways. These may be used in determining drug sensitivity in patients to allow the development of individualized genetic profiles which aid in treating diseases and disorders (e.g. cancer) based on patient response at a molecular level. The present sequence is used in the exemplification of the present invention.

Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;

Query Match 88.7%; Score 1337.2; DB 10; Length 5632;
Best Local Similarity 97.2%; Pred. No. 0;

QY 52 GACAAAGAGCACTAATGCTTTGTGCTGATTCAATTTGAATCGAGGCATTGGAAACCCTG 111

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 112 TATGCCCTTGTTTGTGGAAAGAACCCAGTGACACCATCCTAGCTTCCCTAAAGTTCGAAG 171

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 172 AAGTTAGAGGACTATACACTTTCTTTTGAACTTTTATATAATAAATATTTGCTCTGG-TTTT 230

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 231 GGAACCCAGGACTGTTAGA-GGTGAGTGACAGGCTTAC-AGTGGCCTTAATCCAACCTC 288

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

QY 289 CAGAAATTGCCCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

QY
349 AACTCAATCCCTCTTAATGTACATGGAT-GGCCAAGAGTGATGGCAGCTCTCTTGCCAG 407

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 408 TCCGATGGAGATGGA-GATGCCCTTGTCAATGAAAGGGCCCNCTGTTGTCAATTCGAGCT 466

100

QY 467 ACACAAAGAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526

[illegible]

QY 527 GCAGCCAGACCTTCACACATTCAGAAGACCTTAATAACATGTCTTAATGCAACACCGGC 586

100

QY 587 CTACCCCTCTGTGAACCAAGCAGTCTCTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATA 646

540
580

QY 647 AAAGTCAAGTGC GAACAGAACCTCCCAAGGAAAGAAATTGCAAGGAAATGAATTTAGCT 706

0559

QY 707 GTGAGGTA TGTGGGCAGACATTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAACAC 766

Db 660 GTGAGTATGTGGCGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACAC 719
QY 767 ACAAGATTCTTTTCACTTACGGGTGTAAATGTGCGGAAGAAAGATTCAAGGAGCCTTGGT 826
Db 720 ACAAGATTCTTTTCACTTACGGGTGTAAATGTGCGGAAGAAAGATTCAAGGAGCCTTGGT 779
QY 827 TTCTTAAATACATCGGACRCATTAATGGGAAATCGGGGGCCAGAACAACTGCAGC 886
Db 780 TTCTTAAATACATCGGACRCATTAATGGGAAATCGGGGGCCAGAACAACTGCAGC 839
QY 887 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGCGCGCGCGAGA 946
Db 840 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGCGCGCGCGAGA 899
QY 947 GCATCTCTCTCTCTGCAAAATCTGATGTTTGTGGCTTCTCTATTTCCAAATAAAGAA 1006
Db 900 GCATCTCTCTCTCTGCAAAATCTGATGTTTGTGGCTTCTCTATTTCCAAATAAAGAA 959
QY 1007 GTCTAATTGAGCAGCGCAAGGTGCAACCAAAAAAATCTGTTTGGTATCCAGAGCGCGC 1066
Db 960 GTCTAATTGAGCAGCGCAAGGTGCAACCAAAAAAATCTGTTTGGTATCCAGAGCGCGC 1019
QY 1067 AGACAGACTCTCCACAGGAGGAATGCCGTCTCCGAGGGAGGACTTCTGCGAGTTGTTCA 1126
Db 1020 AGACAGACTCTCCACAGGAGGAATGCCGTCTCCGAGGGAGGACTTCTGCGAGTTGTTCA 1079
QY 1127 ACTTGAGACCAAAATCTCACTCTGAAACGGGGAAGAGCTGTGATGATGCATCCCTCAGC 1186
Db 1080 ACTTGAGACCAAAATCTCACTCTGAAACGGGGAAGAGCTGTGATGATGCATCCCTCAGC 1139
QY 1187 TCGATCCGTTACCACTTCCAGGCTTGCGACTGTGCTACCAAGAAWAGTTGCCATTT 1246
Db 1140 TCGATCCGTTACCACTTCCAGGCTTGCGACTGTGCTACCAAGAAWAGTTGCCATTT 1199
QY 1247 GCCAAGACTGAGGAATTTGGGCAAGAGGAGGACCGCACACGACGATTCGAGTTCCG 1306
Db 1200 GCCAAGACTGAGGAATTTGGGCAAGAGGAGGACCGCACACGACGATTCGAGTTCCG 1259
QY 1307 AGAAGAGCTTGGAGAAACAAATAAGAACCATTTGTGAGGCTCTCGCAAGAGAAAGAGA 1366
Db 1260 AGAAGAGCTTGGAGAAACAAATAAGAGGAGTTGTGAGGCTCTCGCAAGAGAAAGAGA 1319
QY 1367 AGTGCAACACTCCACGCGGAGAGCGCTTCCGTGGAACGGGATCCCAAGTTACCCAGTA 1426
Db 1320 AGTGCAACACTCCACGCGGAGAGCGCTTCCGTGGAACGGGATCCCAAGTTACCCAGTA 1379
QY 1427 GCAAGGAGAGCCACTCACTGCTCCGAGTGGCGCAAGCTTTCAGAACTACCCAGC 1486
Db 1380 GCAAGGAGAGCCACTCACTGCTCCGAGTGGCGCAAGCTTTCAGAACTACCCAGC 1439
QY 1487 TGGTCTTGCACTCCAGGGTCC 1507
Db 1440 TGGTCTTGCACTCCAGGGTCC 1460

RESULT 6

ADR66761

ID ADR66761 standard; DNA; 5632 BP.

XX AC ADR66761;

XX AC ADR66761;

DT 02-DEC-2004 (first entry)

XX Human prostatic carcinoma derived DNA SEQ ID 54 #4.

DE human; cytostatic; diagnosis; prostatic cancer;

KW differential expression analysis; ds.

XX OS Homo sapiens.

XX PN WO2004076614-A2.

XX 10-SEP-2004.

PD

XX 22-FEB-2004; 2004WO-DE000433.
PF 27-FEB-2003; 2003DE-01009985.
XX 14-MAY-2003; 2003DE-01022134.
PR (HINZ/) HINZMANN B.
XX (DAHL/) DAHL E.
PA (ROSE/) ROSENTHAL A.
PA (HERM/) HERMANN K.
PA (PILA/) PILARSKY C.
XX Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
PI Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
PI Xinzhong L, Staub E;
XX WPI; 2004-653386/63.
XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
PT agents.
XX Claim 1; Page 1423-1424; 1607pp; German.
PS This invention describes novel cytosstatic polynucleotide and polypeptide
XX sequences which can be used in a method for diagnosing prostatic cancer
CC or the risk of developing prostatic cancer. Diagnosis is based on
CC determining over transcription or over expression of the sequences in
CC prostatic tissue. Screening for inhibitors of the sequences or detection
CC substances involves a binding assay, any compounds that bind are
CC selected, optionally after deconvolution of mixtures. Detection of a
CC predetermined minimum level of the reporter indicates the presence of
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
CC short-interfering RNA or ribozymes; an organic molecule of molecular
CC weight below 5000, preferably 300, that binds to the polypeptide; an
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human
CC (monoclonal) antibody directed against Ab or any of the above derivatised
CC radioisotope. The polynucleotides are identified in human prostatic
CC cancer by differential expression analysis, using DNA microarrays,
CC between normal and tumorous tissues, with (over)expression being detected
CC by quantitative PCR. Analysis of prostatic cancer samples showed that
CC CD24 was upregulated in many of them. Sections of tissue, isolated from
CC prostatic cancer patients, or subjects at risk, were incubated
CC sequentially with anti-human CD4 murine monoclonal antibodies;
CC biotinylated second antibody; streptavidin-conjugated horseradish
CC peroxidase and then diaminobenzidine as colour former (brown). The
CC samples were counterstained with hemalum (blue). Malignant cells stained
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
CC lymph node metastases were also stained. ADR65805-ADR66954 represent the
CC polynucleotide and polypeptide sequences used in the method of the
CC invention.
XX SQ Sequence 5632 BP; 1637 A; 1167 C; 1257 G; 1571 T; 0 U; 0 Other;

Query Match 88.7%; Score 1337.2; DB 13; Length 5632;

Best Local Similarity 97.2%; Pred. No. 0;

Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 52 GACAAAGAGCACTAATGCTTTGTGCTGATTCATATTGATCGAGCATTTGGAAACCCCTG 111

Db 1 GACAAAGAGCACTAATGCTTTGTGCTGATTCATATTGATCGAGCATTTGGAAACCCCTG 60

QY 112 TATGCTTGTGTTGGAAGAACCAAGTGCACATCATCTGAGCTTCTCTAAAGTTGGAAG 171

Db 61 TATGCTTGTGTTGGAAGAACCAAGTGCACATCATCTGAGCTTCTCTAAAGTTGGAAG 120

QY 172 AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATATAATATTTCTCTGG-TTTT 230

Db 121 AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATATAATATTTCTCTGGTTTTT 180

QY 231 GGAACCCAGGACTGTTAGA-GGGTCAGTGACAGGCTTTAC-AGTGGCCTTAATCCAACTC 288
DB 181 GGAACCCAGGCTGTTAGAGGGGTGAGTGACAAGTCTTTACAAGTGGCCTTATTCCAACTC 240
QY 289 CAGAAATTCGCCCAACGGAATTTTGAGATTATATGCAATCGAAAGTGACAGGAAAATGCC 348
DB 241 CAGAAATTCGCCCAACGGAATTTTGAGATTATATGCAATCGAAAGTGACAGGAAAATGCC 300
QY 349 AACTCAATCCCTCTTAATGTATCATGGAT-GGCCAAGAGTGATGGCAGCTCTCTTGGCAG 407
DB 301 AACTCAATCCCTCTTAATGTATCATGGATGGCAGGAGATGATGGCAGCTCTCTTGGCAG 360
QY 408 TCCGATGAGATGGA-GATGCCTCTGCAATGAAGGGCCNCTGTGTCAATTCGAGCT 466
DB 361 TCCGATGAGATGGAAGATGCCCTGTCAATGAAGGGACCGCTGTGTTCANTCCGAGC 420
QY 467 ACACAAGAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526
DB 421 TACAACAAGAAAAATGTC-ATCCAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT 479
QY 527 GCAGCAGACCTTCACACATTCAGAAGACCTTAATAAATGCTTTAATGCAACACCGGC 586
DB 480 GCAGCAGACCTTCACACATTCAGAAGACCTTAATAAATGCTTTAATGCAACACCGGC 539
QY 587 CTACCCCTCTGTGAACACGAGCTCTTCGGTGTGAAGCAGAGTATCTCAGTCGCTTGATA 646
DB 540 CTACCCCTCTGTGAACACGAGCTCTTCGGTGTGAAGCAGAGTATCTCAGTCGCTTGATA 599
QY 647 AAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAAAATTCGAAGGAAAAATGAATTAGCT 706
DB 600 AAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAAAATTCGAAGGAAAAATGAATTAGCT 659
QY 707 GTGAGGTATGGGAGACACATTTAGATCGCTTTTGTATGTTGAGATCCACATGAGAACAC 766
DB 660 GTGAGGTATGGGAGACACATTTAGATCGCTTTTGTATGTTGAGATCCACATGAGAACAC 719
QY 767 ACAAGATTTCTTCACTTACGGGTCTAACATGTGCGGAAGAGATTCAAGGAGCCTTCGT 826
DB 720 ACAAGATTTCTTCACTTACGGGTCTAACATGTGCGGAAGAGATTCAAGGAGCCTTCGT 779
QY 827 TTCTTAAAAATCACATGCGGACRATAATGGCAAAATCGGGGCCAGAGCAAACTGCGAGC 886
DB 780 TTCTTAAAAATCACATGCGGACACATATGGCAAAATCGGGGCCAGAGCAAACTGCGAGC 839
QY 887 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGTCAGGTGACCGCGGCGAGA 946
DB 840 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGTCAGGTGACCGCGGCGAGA 899
QY 947 GCATCTCTCTCTTGCAGAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAGAAA 1006
DB 900 GCATCTCTCTCTTGCAGAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAGAAA 959
QY 1007 GTCTAATTTGAGCACCGCAAGGTGACACCAAAAAAATGCTTTTCGGTACCAGCGCGC 1066
DB 960 GTCTAATTTGAGCACCGCAAGGTGACACCAAAAAAATGCTTTTCGGTACCAGCGCGC 1019
QY 1067 AGACAGACTCTCCACAGAGGAATGCGGTCTCTCAGGAGGAGACTTCTCGCATGTTGTTCA 1126
DB 1020 AGACAGACTCTCCACAGAGGAATGCGGTCTCTCAGGAGGAGACTTCTCGCATGTTGTTCA 1079
QY 1127 ACTTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCTGTCTCAGATGCAATCCCTCAGC 1186
DB 1080 ACTTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCTGTCTCAGATGCAATCCCTCAGC 1139
QY 1187 TCGATCCGTTACCACTTTCAGGCTTGGCAKCTGGCTACCAAGGAAGAGTTGCCATTT 1246
DB 1140 TCGATCCGTTACCACTTTCAGGCTTGGCACTGGCTACCAAGGAAGAGTTGCCATTT 1199
QY 1247 GCCAAGAGTAGGAATTTGGGCAAGAGGAGCACCGACCAACGACATTCGATTCGG 1306
DB 1200 GCCAAGAGTAGGAATTTGGGCAAGAGGAGCACCGACCAACGACATTCGATTCGG 1259
QY 1307 AGAAGGAGCTTGGAGAAAACAAATAGAACCATTTGTGTGAGGCGCTCTCGCAAGAGAAAGAGA 1366

DB 1260 AGAAGGAGCTTGGAGAAACAATAAGGCACTTGTGAGGCTCTCGCAAGAAAGAGA 1319
QY 1367 AGTGCAAAACACTCCACACGGCAAGCGCCTCCGTGGACGCGGATCCCAAGTTTACCCAGTA 1426
DB 1320 AGTGCAAAACACTCCACACGGCAAGCGCCTCCGTGGACGCGGATCCCAAGTTTACCCAGTA 1379
QY 1427 GCAAGGAGAACCCACTCACTGCTCCGAGTGGGCAAGACTTTCAGAACCTTACCAACCAGC 1486
DB 1380 GCAAGGAGAACCCACTCACTGCTCCGAGTGGGCAAGACTTTCAGAACCTTACCAACCAGC 1439
QY 1487 TGGTCTTCACTCCAGGGTCC 1507
DB 1440 TGGTCTTCACTCCAGGGTCC 1460
RESULT 7
ADR65858
ID ADR65858 standard; DNA; 5632 BP.
XX ADR65858;
XX
DT 02-DEC-2004 (first entry)
XX Human prostatic carcinoma derived DNA SEQ ID 54 #1.
XX human; cytostatic; diagnosis; prostatic cancer;
KW differential expression analysis; ds.
XX Homo sapiens.
XX
FN WO2004076614-A2.
XX
PD 10-SEP-2004.
XX
PF 22-FEB-2004; 2004WO-DE000433.
PR 27-FEB-2003; 2003DE-01009985.
PR 14-MAY-2003; 2003DE-01022134.
XX
(HINZ/) HINZMANN B.
PA (DAHL/) DAHL E.
PA (ROSE/) ROSENTHAL A.
PA (HERM/) HERMANN K.
PA (PILA/) PILARSKY C.
XX
Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
Xinzhong L, Staub E;
XX
WPI; 2004-653386/63.
XX
New nucleic acids, and encoded proteins, from prostatic cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
PT agents.
XX
Claim 1; Page 253-254; 1607pp; German.
XX
This invention describes novel cytostatic polynucleotide and polypeptide
sequences which can be used in a method for diagnosing prostatic cancer
or the risk of developing prostatic cancer. Diagnosis is based on
CC determining over transcription or over expression of the sequences in
CC prostatic tissue. Screening for inhibitors of the sequences or detection
CC substances involves a binding assay, any compounds that bind are
CC selected, optionally after deconvolution of mixtures. Detection of a
CC predetermined minimum level of the reporter indicates the presence of
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
CC short-interfering RNA or ribozymes; an organic molecule of molecular
CC weight below 5000, preferably 300, that binds to the polypeptide; an
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human
CC (monoclonal) antibody directed against Ab or any of the above derivatised
CC with a reporter group, cell toxin, immunostimulatory molecules and/or

CC radioisotope. The polynucleotides are identified in human prostatic
 CC cancer by differential expression analysis, using DNA microarrays,
 CC between normal and tumorous tissues, with (over)expression being detected
 CC by quantitative PCR. Analysis of prostatic cancer samples showed that
 CC CD24 was upregulated in many of them. Sections of tissue, isolated from
 CC prostatic cancer patients, or subjects at risk, were incubated
 CC sequentially with anti-human CD4 murine monoclonal antibodies;
 CC biotinylated second antibody; streptavidin-conjugated horseradish
 CC peroxidase and then diaminobenzidine as colour former (brown). The
 CC samples were counterstained with hemalum (blue). Malignant cells stained
 CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
 CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
 CC lymph node metastases were also stained. ADR65805-ADR66954 represent the
 CC polynucleotide and polypeptide sequences used in the method of the
 CC invention.
 XX
 SQ Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;

Query Match 88.7%; Score 1337.2; DB 13; Length 5632;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 52 GACAAAGACACTAATGCTTGTGCTGATTCTATATTGAATCGAGCATTTGGACCCCTG 111
 DB 1 GACAAAGACACTAATGCTTGTGCTGATTCTATATTGAATCGAGCATTTGGACCCCTG 60
 QY 112 TATGCCCTTGTGTGAAAGAACCACTGACACCATCAGCTTCTTAAAGTTTCGAAG 171
 DB 61 TATGCCCTTGTGTGAAAGAACCACTGACACCATCAGCTTCTTAAAGTTTCGAAG 120
 QY 172 AAGTTAGAGCACTATACATTTCTTTTGAATTTTATATAATATTTCTCTGG-TTTT 230
 DB 121 AAGTTAGAGCACTATACATTTCTTTTGAATTTTATATAATATTTCTCTGGTTTTT 180
 QY 231 GGAACCCAGCACTGTTAGG-GGGTGATGACAGCTCTTAC-AGTGGCTTATCCAACTC 288
 DB 181 GGAACCCAGGCTGTGTAGAGGGGTGATGACAGCTCTTACAGTGGCTTATTTCCAACTC 240
 QY 289 CAGAAATTCGCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348
 DB 241 CAGAAATTCGCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
 QY 349 AACTCAATCCCTTTAATCTAATGAT-GGCCAGAGATGATGGCAGCTCTCTTGGCCAG 407
 DB 301 AACTCAATCCCTTTAATCTAATGATGGCCAGAGATGATGGCAGCTCTCTTGGCCAG 360
 QY 408 TCCGATGGAGTGGG-ATGCTCTTGTCAATGAAGGGCCCTTGTGTCAATCCGAGCT 466
 DB 361 TCCGATGGAGTGGGAGGATGGCTTGTCAATGAAGGGCCCTGTTGTTCATTCGAGC 420
 QY 467 ACACAAAGAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526
 DB 421 TACACAGAAAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT 479
 QY 527 GCAGCCAGACCTTTCACATTCAGAGACCTTAAATAACATGTCTTAAATGCAACACCGGC 586
 DB 480 GCAGCCAGACCTTTCACATTCAGAGACCTTAAATAACATGTCTTAAATGCAACACCGGC 539
 QY 587 CTACCTCTGTGAACACGAGCTTCTTCGGTTGAAGCAGAGATATCTCAGTCCGCTTGATA 646
 DB 540 CTACCTCTGTGAACACGAGCTTCTTCGGTTGAAGCAGAGATATCTCAGTCCGCTTGATA 599
 QY 647 AAAGTCAAGTCGGAACAGAACTCCCAAGGAAAAAGATTGCAAGAAATGAATTTAGCT 706
 DB 600 AAAGTCAAGTCGGAACAGAACTCCCAAGGAAAAAGATTGCAAGAAATGAATTTAGCT 659
 QY 707 GTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGAATTTGAGATFCCAATGAGAACAC 766
 DB 660 GTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGAATTTGAGATFCCAATGAGAACAC 719
 QY 767 ACAAGATTTCTTTCACTTACGGGTGTAAATGTGGGAAAGAGATTCAAGGAGCCTTGGT 826
 DB 720 ACAAGATTTCTTTCACTTACGGGTGTAAATGTGGGAAAGAGATTCAAGGAGCCTTGGT 779

QY 827 TTCTTAAAAATCACATGCGACRCATATGCAAAATCGGGGCGCCAGAACAACTGCAGC 886
 DB 780 TTCTTAAAAATCACATGCGACRCATATGCAAAATCGGGGCGCCAGAACAACTGCAGC 839
 QY 887 AAGGCTTGGAGAGTAGTCCAGCAACGATCAACAGAGTCTCCAGGTGACGGGGCCGAGA 946
 DB 840 AAGGCTTGGAGAGTAGTCCAGCAACGATCAACAGAGTCTCCAGGTGACGGGGCCGAGA 899
 QY 947 GCATCTCTCTCTCTTTCGCAAAATCTCATGGTTTGTGGCTTCTTATTTCCAAATAAAGAA 1006
 DB 900 GCATCTCTCTCTCTTTCGCAAAATCTCATGGTTTGTGGCTTCTTATTTCCAAATAAAGAA 959
 QY 1007 GTCTAATTGAGCACCGCAGGTGCACACCAAAAAAATCTCTTTTCGGTACAGCAGCGCGC 1066
 DB 960 GTCTAATTGAGCACCGCAGGTGCACACCAAAAAAATCTCTTTTCGGTACAGCAGCGCGC 1019
 QY 1067 AGACAGACTCTCCACAGGAGGAATGCCGTCTCGAGGGAGGACTTCTCGAGTTGTTCA 1126
 DB 1020 AGACAGACTCTCCACAGGAGGAATGCCGTCTCGAGGGAGGACTTCTCGAGTTGTTCA 1079
 QY 1127 ACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAGCCTGTGATGCATCCCTCAGC 1186
 DB 1080 ACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAGCCTGTGATGCATCCCTCAGC 1139
 QY 1187 TCGATCCGTTTCAACCACTTTCAGAGGCTTGGCACTTGGCACTTGGCACTTGGCACTTT 1246
 DB 1140 TCGATCCGTTTCAACCACTTTCAGAGGCTTGGCACTTGGCACTTGGCACTTT 1199
 QY 1247 GCCAAGAGTGAAGAAATTTGGGGCAAGAGGAGCAACGACGATTCGAGTTCCG 1306
 DB 1200 GCCAAGAGTGAAGAAATCGGGGCAAGAGGAGCAACGACGATTCGAGTTCCG 1259
 QY 1307 AGAAGGAGCTTGGAGAAACAAATAGAACCTTGTGAGGCTCTCGAAGAGAAAGAGA 1366
 DB 1260 AGAAGGAGCTTGGAGAAACAAATAGAACCTTGTGAGGCTCTCGAAGAGAAAGAGA 1319
 QY 1367 AGTGAACACACTCCCAACGCGGAGGCGCTTCCGTGGAACGCGGATCCCAAGTTACCCAGTA 1426
 DB 1320 AGTGAACACACTCCCAACGCGGAGGCGCTTCCGTGGAACGCGGATCCCAAGTTACCCAGTA 1379
 QY 1427 GCAAGGAGAGCCCACTCACTGCTCCGAGTGGGGCAAGAGCTTTCAGAACTACCCAGCAG 1486
 DB 1380 GCAAGGAGAGCCCACTCACTGCTCCGAGTGGGGCAAGAGCTTTCAGAACTACCCAGCAG 1439
 QY 1487 TGGTCTTGCACCTCCAGGGTCC 1507
 DB 1440 TGGTCTTGCACCTCCAGGGTCC 1460

RESULT 8
 ADP07267
 ID ADP07267 standard; DNA; 5653 BP.
 XX
 AC ADP07267;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human ZNF217 DNA.
 XX
 KW ds; proliferative disease; breast; methylation; CpG; bisulfite; human.
 OS Homo sapiens.
 XX
 PN DE10255104-Al.
 XX
 PD 11-MAR-2004.
 XX
 PF 26-NOV-2002; 2002DE-01055104.
 XX
 PR 27-AUG-2002; 2002DE-01039313.
 XX
 PA (EPIC-) EPIGENOMICS AG.

XX
PI Mater S;
XX
DR WPI; 2004-284340/27.
DR GENBANK; NM_006526.
XX
PT Analyzing proliferative diseases of breast cells, useful e.g. for
PT diagnosis, prognosis and treatment of breast cancer, by determining
PT methylation status of specific genes.
XX
PS Claim 1; Page; 22pp; German.
XX
CC This invention describes a novel method of analysing proliferative
CC diseases of breast cells by determining the methylation status of certain
CC genes. The invention also describes nucleic acids or their complements,
CC oligomers, especially oligonucleotides or peptide nucleic acid oligomers,
CC that hybridise to, or are identical with, any of the nucleic acids, the
CC preparation of an oligomer array for analysing proliferative diseases of
CC breast cells that are associated with the methylation status of CpG
CC dinucleotides of the genes by bonding at least one oligomer to a solid
CC phase and a kit comprising a bisulfite reagent and the oligomer. The
CC nucleic acids are genomic sequences (5' - and/or regulatory and/or CpG-
CC rich regions). The base sequence of the oligomer includes at least one
CC CpG island, especially with C in the middle third of the sequence. The
CC process involves a genomic DNA sample treated chemically, specifically
CC with a bisulfite reagent, to convert non-methylated C to uracil or some
CC other base having base-pairing properties different from those of C.
CC Fragments of the treated DNA are then amplified, using primers and a
CC polymerase and the methylation status of the genomic CpG dinucleotides is
CC determined by analysis of the amplicons, particularly by hybridisation to
CC the oligomer, optionally with extension of the hybridised oligomer by at
CC least one base, or detection is by sequencing. The amplification may use
CC methylation-specific primers. Alternatively, genomic DNA is extracted
CC from a sample and digested with methylation-specific restriction enzymes,
CC then the digestion fragments detected, optionally after amplification. In
CC either method, more than 10 fragments of 100-200 bp are amplified in a
CC single reaction vessel, using a heat-resistant DNA polymerase in PCR. The
CC amplicons carry detectable markers, e.g. fluorophores, radioisotopes
CC and/or releasable fragments of known mass that can be detected by mass
CC spectrometry. The method is used for characterisation, classification,
CC differentiation, staging, phase-estimation, diagnosis and/or therapy of
CC proliferative diseases of breast cells. The method provides very specific
CC classification of proliferative diseases, allowing better treatment. It
CC can both characterise methylation status and detect single-nucleotide
CC polymorphisms. This sequence represents human gene used to illustrate the
CC method of the invention. NOTE: This sequence does not appear in the
CC printed specification but has been retrieved from Genbank.
XX
SQ Sequence 5653 BP; 1656 A; 1167 C; 1258 G; 1572 T; 0 U; 0 Other;

QY	349	AACCTCAATCCCTCTTAATGTATCATGGAT- GGCCAAGAGTGATGGCAGCTCTCTTGGCAG	407
DB	301	AACCTCAATCCCTCTTAATGTATCATGGATGGGCGAAGAGTGATGGCAGCTCTCTTGGCAG	360
QY	408	TCCGATGAGATGGA- GATGCTTGTCAATGAAAGGGCCNCTGTGTGTCATATTCGAGCT	466
DB	361	TCCGATGAGATGAGGATGCTTGTCAATGAAAGGACCGCTGTGTTCATTCGAGC	420
QY	467	ACACAAAGAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTCT	526
DB	421	TACACAGAAAAAATGTC- ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTCT	479
QY	527	GCAGCCAGACCTTCACACATTCAGAGACCTTATTAACATCTCTTAATGCAACACCGC	586
DB	480	GCAGCCAGACCTTCACACATTCAGAGACCTTATTAACATCTCTTAATGCAACACCGC	539
QY	587	CTACCTCTGTGAACCCAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCCGCTTGATA	646
DB	540	CTACCTCTGTGAACCCAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCCGCTTGATA	599
QY	647	AAAGTCAAGTGCAGAACAACTCCCAAGGAAAAAGATTGCAAGAAAAATGAATTTAGCT	706
DB	600	AAAGTCAAGTGCAGAACAACTCCCAAGGAAAAAGATTGCAAGAAAAATGAATTTAGCT	659
QY	707	GTGAGTATGTGGGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAAACAC	766
DB	660	GTGAGTATGTGGGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAAACAC	719
QY	767	ACAAAGATTCTTTTCACTTACCGGTGTAAACATGTGGGAAGAAAGATTCAAGGAGCCTTGT	826
DB	720	ACAAAGATTCTTTTCACTTACCGGTGTAAACATGTGGGAAGAAAGATTCAAGGAGCCTTGT	779
QY	827	TTCTTAAAAATCAATGCGGACACRCAATATGGCAATCGGGGCGCAGAGCAAACTGAGC	886
DB	780	TTCTTAAAAATCAATGCGGACACACATAATGGCAATCGGGGCGCAGAGCAAACTGAGC	839
QY	887	AAGCTTGGAGAGTAGTCCACACACGATCAACAGGTGCTCCAGGTGACCGCGCGGAG	946
DB	840	AAGCTTGGAGAGTAGTCCACACACGATCAACAGGTGCTCCAGGTGACCGCGCGGAG	899
QY	947	GCATCTCTCTCTTTCGCAAAATCTGCATGGTGTGGTCTCTCTATTTCCAAATAAGAAA	1006
DB	900	GCATCTCTCTCTTTCGCAAAATCTGCATGGTGTGGTCTCTCTATTTCCAAATAAGAAA	959
QY	1007	GTCTAATTTAGGACCGCAAGGTGACACCAAAAAAATCTGTTTCCGTTACCGACGCGGC	1066
DB	960	GTCTAATTTAGGACCGCAAGGTGACACCAAAAAAATCTGTTTCCGTTACCGACGCGGC	1019
QY	1067	AGACAGACTCTCCCAAGGAGGAAATGCCGTCTCCGAGGGAGGACCTTCTGCAAGTTGTTCA	1126
DB	1020	AGACAGACTCTCCCAAGGAGGAAATGCCGTCTCCGAGGGAGGACCTTCTGCAAGTTGTTCA	1079
QY	1127	ACTTGAGACCAAAATCTCACCTCGAAACGGGGAAGAACCTGTGATGATGCATCCCTCAGC	1186
DB	1080	ACTTGAGACCAAAATCTCACCTCGAAACGGGGAAGAACCTGTGATGATGCATCCCTCAGC	1139
QY	1187	TCGATCCCTTCCACCTTTCAGGGCTTGCGACTGGCTACCAAGGAAAGTTGGCATTT	1246
DB	1140	TCGATCCCTTCCACCTTTCAGGGCTTGCGACTGGCTACCAAGGAAAGTTGGCATTT	1199
QY	1247	GCCAAGAGTCAAGGAATTTGGGGAAGAGGAGGACCGAACACCGATTCGAGTTCCG	1306
DB	1200	GCCAAGAGTCAAGGAATTTGGGGAAGAGGAGGAGGACCGAACACCGATTCGAGTTCCG	1259
QY	1307	AGAAGGACTTTGGGAAAAAATAATAGAACCAATTTGTGAGGCGCTCTCGAAGAGAAAGAGA	1366
DB	1260	AGAAGGACTTTGGGAAAAAATAATAGAGGAGTTGTGAGGCGCTCTCGAAGAGAAAGAGA	1319
QY	1367	AGTGCAAAACATCTCCACCGGAGAGGCGCTCTCGTGGAAGCGGATCCCAAGTTACCCAGTA	1426
DB	1320	AGTGCAAAACATCTCCACCGGAGAGGCGCTCTCGTGGAAGCGGATCCCAAGTTACCCAGTA	1379

QY 1427 GCAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTCAGAACCTACCACGAC 1486
DB 1380 GCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTCAGAACCTACCACGAC 1439
QY 1487 TGGCTTTGCACTCCAGGGTCC 1507
DB 1440 TGGTCTTGCACTCCAGGGTCC 1460

RESULT 9
ACN44987
ID ACN44987 standard; cDNA; 5252 BP.
XX
AC ACN44987;
XX
DT 18-NOV-2004 (first entry)
DE Human mRNA sequence hCT1950762.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
PI Morris DW;
XX
XX WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1709; Opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA gene
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 5252 BP; 1507 A; 1101 C; 1177 G; 1467 T; 0 U; 0 Other;

Query Match 75.2%; Score 1132.6; DB 11; Length 5252;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1180; Conservative 3; Mismatches 28; Indels 3; Gaps 3;
QY 296 TGCCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCCAATCAA 355
DB 1 TGCCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCCAATCAA 60
QY 356 TCCCTCTTAATGTACATGGAT-GGCCAAGAGTGATTGGCAGCTCTCTTGCCAGTCCGATG 414
DB 61 TCCCTCTTAATGTACATGGATGGCGGAGGATGGCAGCTCTCTTGCCAGTCCGATG 120
QY 415 GAGATGGA-GATGCCCTTGTCAATGAAGGGCCCNCTGTTGTCATTCGAGCTACACAA 473
|||||

DB 121 GAGATGGAGGTGCGCTTGTCAATGAAGGACGCGCTGTTGTTCCATTCCGAGCTACACAA 180
QY 474 GAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTTGGATTGCATGTTCTGACGCCA 533
DB 181 GAAAAAATGTGTC-ATCCAATCGAGGGGTATATGCCCTTTGGATTGCATGTTCTGACGCCA 239
QY 534 GACCTTTCACACATTTCAGAAAGACCTTAAATAACATGTCTTAATGAACAACGCGGCTACCT 593
DB 240 GACCTTTCACACATTTCAGAAAGACCTTAAATAACATGTCTTAATGAACAACGCGGCTACCT 299
QY 594 CTGTGAACCCAGCAGTTCTTTCGGGTTGAACGACAGATATCTCAGTCCGCTTGTATAAAGTCA 653
DB 300 CTGTGAACCCAGCAGTTCTTTCGGGTTGAACGACAGATATCTCAGTCCGCTTGTATAAAGTCA 359
QY 654 AGTGGCAACAGAAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGT 713
DB 360 AGTGGCAACAGAAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGT 419
QY 714 ATGTGGGACAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATGAGAACACACAAGA 773
DB 420 ATGTGGGACAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATGAGAACACACAAGA 479
QY 774 TTCTTTTCACTTACGGGTGTAAACATGTGCGGAAGAAAGATTTCAAGGAGCCTTGTGTTCTTAA 833
DB 480 TTCTTTTCACTTACGGGTGTAAACATGTGCGGAAGAAAGATTTCAAGGAGCCTTGTGTTCTTAA 539
QY 834 AAATCACATGCGGACRCATAATATGGCAAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTT 893
DB 540 AAATCACATGCGGACRCATAATATGGCAAAATCGGGGGCCAGAAGCAAACTGCAGCAAGGCTT 599
QY 894 GGAGAGTAGTCCAGCAACGATCAAGAGGTGCTCAGGTGCGACGGCGCGAGAGCATCTC 953
DB 600 GGAGAGTAGTCCAGCAACGATCAAGAGGTGCTCAGGTGCGACGGCGCGAGAGCATCTC 659
QY 954 CTCTCTTGCRAAAATCTGCATGCTTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAAT 1013
DB 660 CTCTCTTGCRAAAATCTGCATGCTTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAAT 719
QY 1014 TGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTTCGGTACCAGCAGCGCGCAGACAGA 1073
DB 720 TGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTTCGGTACCAGCAGCGCGCAGACAGA 779
QY 1074 CTCTCCACAGAGAGAAATGCCCTCTCGAGGAGGAGACTTCTTCGAGTTGTTCAACTTGAG 1133
DB 780 CTCTCCACAGAGAGAAATGCCCTCTCGAGGAGGAGACTTCTTCGAGTTGTTCAACTTGAG 839
QY 1134 ACCAAATCTCACCTGTAACCGGGGAAGGAGCTGTTCAGATGCATCCCTCAGCTCGATCC 1193
DB 840 ACCAAATCTCACCTGTAACCGGGGAAGGAGCTGTTCAGATGCATCCCTCAGCTCGATCC 899
QY 1194 GTTCAACCACTTCCAGGCTTGGCACTGTGCTACCAAGGAAGWAGTTGCCAATTTGCCAAGA 1253
DB 900 GTTCAACCACTTCCAGGCTTGGCAGCTGGCTACCAAGGAAGAAAGTTGCCAATTTGCCAAGA 959
QY 1254 AGTGAAGGAATTTGGGGCAAGAGGAGGACACCGACAAACGACGATTCGAGTTCCGAGAAGGA 1313
DB 960 AGTGAAGGAATTCGGGGCAAGAGGAGGACACCGACAAACGACGATTCGAGTTCCGAGAAGGA 1019
QY 1314 GCTTCGGAAGAACAAATAGAACCATTGTGCGAGCCTCTCGCAAGAGAGAGAGAGTGCA 1373
DB 1020 GCTTCGGAAGAACAAATAGAAGGAGTTGTGCGAGCCTCTCGCAAGGAAGAGAGTGCA 1079
QY 1374 ACACCTCCACGCGGAGCGCCCTCCGCTGGACCGGATCCCAAGTTACCAGTAGCAAGGA 1433
DB 1080 ACACCTCCACGCGGAGCGCCCTCCGCTGGACCGGATCCCAAGTTACCAGTAGCAAGGA 1139
QY 1434 GAAGCCCACTCACTGCTCGAGTGGCGGCAAGAGCTTTTCAGAACTTACCACAGCTGTGCTTT 1493
DB 1140 GAAGCCCACTCACTGCTCGAGTGGCGGCAAGAGCTTTTCAGAACTTACCACAGCTGTGCTTT 1199
QY 1494 GCACCTCCAGGGTCC 1507
DB 1200 GCACCTCCAGGGTCC 1213

RESULT 10
AAV09024
ID AAV09024 standard; DNA; 3183 BP.
AC AAV09024;
XX
XX
XX 21-JUL-1998 (first entry)
XX
XX Homo sapiens 20q13 amplicon ZABC-1 cDNA sequence.
XX
XX 20q13 amplicon; chromosome 20; tumour; detection; ZABC-1 gene;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentata;
KW Leber's congenital amaurosis; zinc finger amplified in breast cancer; ds.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..3183
FT /*tag= a
FT /product= "ZABC1 protein"
XX
XX WO9802539-A1.
XX
XX 22-JAN-1998.
XX
XX 15-JUL-1997; 97WO-US012343.
XX
XX 15-JUL-1996; 96US-00680395.
PR 16-OCT-1996; 96US-00731499.
PR 17-JAN-1997; 97US-00785532.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX WPI; 1998-110587/10.
DR P-FSDB; AAW23975.
XX
XX New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
PT treating diseases.
XX
XX Claim 1; Page 67-69; 9lpp; English.
XX
XX The sequence is that of the cDNA sequence encoding ZABC-1 (zinc finger
CC amplified in breast cancer). It maps to the core of the 20q13.2 amplicon
CC and is overexpressed in primary tumours and breast cancer cell lines
CC having 20q13.2 amplification. The sequence can be used as a probe for the
CC detection of chromosomal abnormalities at 20q13. It and other sequences
CC isolated from the 20q13 amplicon are consistently amplified in primary
CC tumours. These sequences are useful as probes or as probe targets for
CC monitoring the relative copy number of corresponding sequences from a
CC biological sample such as tumour cells. The sequences can also be used in
CC therapeutic applications for modulating the expression of the endogenous
CC gene or the activity of the gene product. Examples of therapeutic
CC approaches include antisense inhibition of gene expression, gene therapy,
CC and monoclonal antibodies that specifically bind the gene products. The
CC products can also be used in the treatment of other diseases, e.g. age-
CC related macular degeneration, Leber's congenital amaurosis and retinitis
CC pigmentata
XX
SQ Sequence 3183 BP; 925 A; 796 C; 758 G; 698 T; 0 U; 6 Other;
Query Match 72.1%; Score 1087.2; DB 2; Length 3183;
Best Local Similarity 95.7%; Pred No. 0;
Matches 1138; Conservative 9; Mismatches 38; Indels 4; Gaps 3;
320 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTTTAATGTACATGATGCG 379
1 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTTTAATGTACATGATGCG 60

QY 380 CAAGAGTCATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGA-CATGCTTGTCAATGA 438
DB 61 CCAGAAAGTGAATG--GCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATGA 118
QY 439 AAGGCCCCNCTGTGTCAATTCGAGCTACACAAAGAAAAAATGTCATCAATCGAATCGAG 498
DB 119 AAGGACCGCTGTGTTCATTCGAGCTACACAAAGAAAAAATGTC-ATCCAAATCGAG 177
QY 499 GGGAAATATGCCCTTGGATTGCAATGTTCTGAGCCAGACCTTCAACATTCAGAAACCTT 558
DB 178 GCGTATATGCCCTTGGATTGCAATGTTCTGAGCCAGACCTTCAACATTCAGAAACCTT 237
QY 559 AATAAATATGTTAATGCAACACCGGCTTACCTCTGTAACAGCAGATGTTCTTGGGTT 618
DB 238 AATAAATATGTTAATGCAACACCGGCTTACCTCTGTAACAGCAGATGTTCTTGGGTT 297
QY 619 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCAGAACAGAACCTCCCAAGAA 678
DB 298 GAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGCAGAACAGAACCTCCCAAGAA 357
QY 679 AAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGATCGCT 738
DB 358 AAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGATCGCT 417
QY 739 TTTGATGTTGAGATCCCATGAGACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 798
DB 418 TTTGATGTTGAGATCCCATGAGACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 477
QY 799 TCGGAGAAAGATTCAGGAGCTTTGGTTTCTTAAAAATCACATCGGACRACATATGCG 858
DB 478 TCGGAGAAAGTTSRRSAGCTTGGTTTCTTAAAAATCACATCGGACRACATATGCG 537
QY 859 AAATCGGGGGCAGAAAGCAAACTGACAGAGCTTTGGAGAGTAGTCCAGAACAGATCAAC 918
DB 538 AAATCGGGGGCAGAAAGCAAACTGACAGAGCTTTGGAGAGTAGTCCAGAACAGATCAAC 597
QY 919 GAGTCTGTCAGGTGACGCGGCGAGAGCATCTCTCTCTTGCAGAAATCTGCAATGTT 978
DB 598 GAGTCTGTCAGGTGACGCGGCGAGAGCATCTCTCTCTTGCAGAAATCTGCAATGTT 657
QY 979 TGTGCTCTCTATTTCCAAATAAAGAAAGTCTAATTGAGACCCGCAAGGTGCACACCAA 1038
DB 658 TGTGCTCTCTATTTCCAAATAAAGAAAGTCTAATTGAGACCCGCAAGGTGCACACCAA 717
QY 1039 AAACTGCTTTTCGGTACAGAGCGCGCAGACAGACTCTCCACAGAGAGGAATGCGCTCC 1098
DB 718 AAACTGCTTTTCGGTACAGAGCGCGCAGACAGACTCTCCACAGAGAGGAATGCGCTCC 777
QY 1099 TCGAGGAGGAGCTTCTCGAGTTGTTCACTTGAGACCAAAATCTCACCTGAAACGGG 1158
DB 778 TCGAGGAGGAGCTTCTCGAGTTGTTCACTTGAGACCAAAATCTCACCTGAAACGGG 837
QY 1159 AAGAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTTCAACACCTTCAGGCTTGGCAK 1218
DB 838 AAGAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTTCAACACCTTCAGGCTTGGCAG 897
QY 1219 CTGGCTACCAAGGAAGTTGTCATTTGCCAAGAGTGAAGGAATTTGGGCAAGAGGG 1278
DB 898 CTGGCTACCAAGGAAGTTGTCATTTGCCAAGAGTGAAGGAATTCGGGGCAAGAGGG 957
QY 1279 AGCACCAGAACGAGATTTCGAGTTCCGAGAGAGGAGCTTGGAGAAACAAATAAGAACCAT 1338
DB 958 AGCACCAGAACGAGATTTCGAGTTCCGAGAGAGGAGCTTGGAGAAACAAATAAGGGCAGT 1017
QY 1339 TGTGAGGCTCTCGCAAGAGAAAGAGAGTGCACAACTCCACGCGGAGAGCGCCCTCC 1398
DB 1018 TGTGAGGCTCTCGCAAGAGAAAGAGAGTGCACAACTCCACGCGGAGAGCGCCCTCC 1077
QY 1399 GTGAGACGCGGATTCCTCAAGTTTACCGAGTAGCAGAGAGAGCCCACTCCTCTCCAGTGC 1458
DB 1078 GTGAGACGCGGATTCCTCAAGTTTACCGAGTAGCAGAGAGAGCCCACTCCTCTCCAGTGC 1137

QY 1459 GGCAAGCTTTTCAGAACCTTACCAACGAGCTGGTCTTTCGACATCCAGGGTCC 1507
 |||||
 Db 1138 GGCAAGCTTTTCAGAACCTTACCAACGAGCTGGTCTTTCGACATCCAGGGTCC 1186
 |||||

RESULT 11

ACN44984

ID ACN44984 standard; DNA; 26345 BP.

XX

AC

ACN44984;

XX

DT

18-NOV-2004 (first entry)

XX

DE

Mouse genomic sequence mCG6549.

XX

KW

Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.

XX

OS

Mus musculus.

XX

PN

WO2003073826-A2.

XX

PD

12-SEP-2003.

XX

PP

28-FEB-2003; 2003WO-US006235.

XX

PR

01-MAR-2002; 2002US-00087192.

XX

PA

(SAGR-) SAGRES DISCOVERY.

XX

PI

Morris DW;

XX

DR

WPI; 2003-328604/31.

XX

PT

Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

XX

PT

comprises a nucleotide sequence.

XX

PS

Claim 1; SEQ ID NO 1705; Opp; English.

XX

CC

The present invention relates to novel DNA and protein sequences which

CC

are associated with carcinomas. The sequences are useful for: (i) for

CC

screening drug candidates; (ii) for screening of bioactive agent capable

CC

of binding to Carcino Associated Protein (CAP); (iii) for screening of

CC

a bioactive agent capable of modulating the activity of CAP; (iv) for

CC

evaluating the effect of a candidate carcinoma drug; (v) for diagnosing

CC

carcinoma; (vi) for inhibiting the activity of CAP; (vii) as a biochip;

CC

carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;

CC

(x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for

CC

determining Carcino Associated (CA) gene copy number. In addition, the

CC

CA genes are useful as DNA vaccines and the CAP are useful as markers of

CC

carcinoma including lymphoma. The present sequence is one such CA coding

CC

sequence. Note: This patent is an equivalent to basic patent

CC

US2002182586A1, for which no sequence data was published

CC

XX

SQ

Sequence 26345 BP; 6286 A; 5855 C; 6718 G; 7486 T; 0 U; 0 Other;

XX

Query Match

Best Local Similarity 38.1%; Score 573.6; DB 11; Length 26345;

Matches 1028; Conservative 3; Mismatches 448; Indels 88; Gaps 10;

XX

QY

2 CAGGTTGCGGATGACCTTCTTGTCTCAATTTGAACACATCTTCAATGAGACAAAGAGC 61

Db

9674 CACATGCGGGATCAACGCTCTGTCTCACTGAGCCACTCACAGTGGAGATGGACAGAG 9733

QY

62 ACTAATGCTTTGGCTGATTCATATTTGAATCGAGGCATTGGGACCCCTGTATGCTTGT 121

Db

9734 CTGTTGCCCTCGTGTGACGACATCTGACTCACAGCCCTGGGGCAGCTTTGGATGCTTGT 9793

QY

122 TTGTGGAAGAACCACTGACACCATCACTGAGCTTCTCTAAAGTTCGAAAGATTAGAGG 181

Db

9794 TTGTGCAAGAGCCA---TCTCCATCAGCTCTCCCAAGCTCC-----GG 9838

QY

182 ACTATACACTTTCTTTTGAACCTTTTATAATAATATTTGCTCTGGTFTTTGGAACCCAGGA 241

Db

|||||

Db 9839 AATATACTGTCCTCTGACCTT-----AGAGGTGCTCTGTATTTTGAACCCAGGG 9889
 QY 242 CTGTTAGAGGGTGAG---TGACAGGTCTTACAGTGGCCTTAATCCAACTCCAGAAATTGC 298
 Db 9890 CTCTTAGAGGTTCTGAGTGGTGATATCTCACAGGCGGCTTATTTCTAAGTGCAGAGATCC 9949
 QY 299 CCAACGGAACTTTGAGATTATATGCAATCGA-----AAGTCACAGGAACAACTGCCAACT 352
 Db 9950 TCAGCAGATAGTGGATTACCCATCTCTGTACAGACCCCTGAGTGGTGGCAGCATGCCGACC 10009
 QY 353 CAATCCCTCTTAATCTACATGGA-TGGCCAAAGAGTATGGCAGCTCTCTTCCAGTCCG 411
 Db 10010 CAGTCCCTCTCTGTATACATGACGCGGCGGAAGTCTCTCAGCAGCTCTCTAGCTCCCAAG 10069
 QY 412 ATGGAGATGGAGATGCCCTTGTCAATCAAAAGGCCCNCTGTTGTCTCAATTCGAGCTACACA 471
 Db 10070 ATGGAGTGGATGATGTGTGCCATAAAGGGCGGTGGCAGTCCCTTCCGAGCTGCT 10129
 QY 472 AAGAAAAAATGTCATTCGAAATCGAGGGAATATGCCCTTGGATTCGATTCGTCAGC 531
 Db 10130 CAGGAGAAAGCATGGCCGTGGCAGAGGGCCACATGCCCTCGATTCGATTCGTCAGC 10189
 QY 532 CAGACCTTCACACATTCAGAGACCTTAAATAACATGCTTAAATGCACACCGGCTTACC 591
 Db 10190 CAGGCTCTCTCTCAGGCGGAGGATCTCAGTCAGCAGCTGCTCTGAGCACCAGGCCACC 10249
 QY 592 CTCTGTGAACACAGCAGTCTCTCGGCTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAAGT 651
 Db 10250 CTCTCGGAGCCAGCTGCTCTGCTGTGGAGGCGGAGTACCTAAGTCCCTTGTATAAAGCT 10309
 QY 652 CAAAGTGCAGAACAGAACCTCCCAAGAAAAAGAAATGCAAGGA---AAATGAATTTAGCTGT 708
 Db 10310 CTGGAGCCAAACAGAGCCAGCATTTGGAGAAGAGTGGCGAAGACCCCGAGGAGTTGAGCTGT 10369
 QY 709 GAGGTATGTGGCAGACACATTTAGACTCCGCTTTTGTGATTTGAGATCCACATGAGAACACAC 768
 Db 10370 GATGTGTGGGCAGACATTTCCAGTGGCTTTTGTGATTTGAGAGGCCACATGAAGAAGCAT 10429
 QY 769 AAGAGATCTTTTACATTTAGCGGTGTAAACATGTGCGGAAGAAAGATTCAAGAGAGCTTGGTTT 828
 Db 10430 AAGGACTCTTTCACGTATGGGTGCGAGCATGTGCGGAGGAGGATTTCAAGAGGCGGTGTTT 10489
 QY 829 CTTAAAAATCAGATCGGACACATTAATGGCAATCGGGGGCCAGAAAGCAAACTGCAGCAA 888
 Db 10490 CTGAAGAACCAATCGGACACACAAATGCAAGTCTGGCACCAGGAGCAAGCTTCAGCAA 10549
 QY 889 GCTTGGAGAGTAGTCCAGCAACGATCAACAGGTCTGTCAGGTGCACGCGGCCGAGAGC 948
 Db 10550 GGCATGGAG---AGTCCAGTCAACATCAATGAAGTGGTCCAGCCGACGCCCCCTGGGAGC 10606
 QY 949 ATCTCTCTCTCTTGCAAAATCTGCAATGTTGTGGCTTCTCTATTTTCCAAATAAAGAAAGT 1008
 Db 10607 ATCTCAGCCCTACAGATCTGCAATGTTGTGGGCTTCTCTTCCCAANTAAAGCAGAGC 10666
 QY 1009 CTAATTGAGACCCGCAAGGTGCACACCAAAAAAACTGCTTTTCGTAACAGCAGCGCGAG 1068
 Db 10667 CTCATTGAGCAGCAAGGTTCAAGCAAGAACTGTCTCCAGTGCAGCAACGTTGTC 10726
 QY 1069 ACAGACTCTCCACAGGAGGAATGCCGCTCTCGAGGAGGAGCTTCTGAGTGTGTCAC 1128
 Db 10727 CCGTGTGACCAACCGAGAGAAACCCACGCTCCCGAGGGAAGTGTGTCAGTTTTTGAAC 10786
 QY 1129 TTGAGACCAAAATCTCACCTGAAACGGGAAGAGCCCTGTGAGATGCATCCCTCAGCTC 1188
 Db 10787 TTGAGACCAAGATCAATGTCAGGTAGTACAGTGAAGCCCATGACCTGCATACCTAGCTT 10846
 QY 1189 GATCGCTTTCACACCTTCCAGCTTTGGCAKCTGGCTTACCAAGAAAGTATGCCATTTTC 1248
 Db 10847 GACCCGTTTCCACCTTACCAGGATGGCAGTGGCTTACCAAGAAAGTATGCCGCTTTC 10906
 QY 1249 C---AAGAAAGTGAAGAAATTTGGGGCAAGAGGGAGCAGCAACGATTCGAGTTCC 1305
 Db 10907 CAGGAAGAGGTGAAGAGTCAAGCCCAAGAGGAAGCAGACAATGACGACTCATGCTCA 10966

QY 1306 GAGAGGAGCTGGAG----- 1321
Db 10967 GAGAAAGAGGAACTAGGGGAAATATGGGTGGGGGTAGGGGGAAGGGTCTGGAAAGTCC 11026
QY 1322 -AAACAAATAAGAACCATTTGTGCAGGCTCTCGCAAGAGAAAGAGAGTGCACACATCC 1380
Db 11027 AAAACAAGTAAAGACAGTTGTCCAGGTCTCTCCCAAGACAAAGAGAGCCTAGACATGCT 11086
QY 1381 CACGGCGAAGCGCCCTCCGTGACCGGGATCCCAAGTTCACCAAGTAGCAGAGAGAGAGGCC 1440
Db 11087 AATAGTGAAGTCCCTTCTGGGGATAGTACCCCAAGTTGTCCAGTAGCAAGAGAGAGGCC 11146
QY 1441 ACTCACTGCTCGAGTGGGCAAGCTTTCAGAACTACCAAGCTGCTTTCAGTCC 1500
Db 11147 ACGCACTGTTGTAGTGCAGCAAGCCTTCAGGACATACCAAGCTGCTGCTGCACTCG 11206
QY 1501 AGGGTCC 1507
Db 11207 AGGGTGC 11213

RESULT 12

ACN44985
ID ACN44985 standard; cDNA; 3016 BP.

XX AC ACN44985;

XX XX 18-NOV-2004 (first entry)

XX DE Mouse mRNA sequence MCT6076.

XX KW Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ds.

XX OS Mus musculus.

XX PN WO2003073826-A2.

XX PD 12-SEP-2003.

XX PF 28-FEB-2003; 2003WO-US006235.

XX PR 01-MAR-2002; 2002US-00087192.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW;

XX WPI; 2003-328604/31.

XX PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
comprises a nucleotide sequence.

XX PS Claim 1; SEQ ID NO 1706; Opp; English.

XX CC The present invention relates to novel DNA and protein sequences which
are associated with carcinomas. The sequences are useful for: (i) for
screening drug candidates; (ii) for screening of bioactive agent capable
of binding to Carcino Associated Protein (CAP); (iii) for screening of
a bioactive agent capable of modulating the activity of CAP; (iv) for
evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
(x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
determining Carcino Associated (CA) gene copy number. In addition, the
CA genes are useful as DNA vaccines and the CAP are useful as markers of
carcinoma including lymphoma. The present sequence is one such CA gene
sequence. Note: This patent is an equivalent to basic patent
US2002182586A1, for which no sequence data was published

XX SQ Sequence 3016 BP; 771 A; 838 C; 847 G; 560 T; 0 U; 0 Other;

XX Query Match 35.6%; Score 537; DB 11; Length 3016;

Best Local Similarity 68.6%; Pred. No. 2.6e-163;
Matches 834; Conservative 3; Mismatches 324; Indels 55; Gaps 5;

QY 344 ATGCCAACTCAATCTCTTAATGTATCATGGA-TGGCCAAGAGTATGTCGACGTCTCTTT 402
Db 1 ATGCCGACCAAGTCTCTCTGTGTACATGGACGGCGGGAAGTCTCTCAGCAGCTCTCTA 60
QY 403 GCCAGTCCGATGGAGATGCCCTTGTCAATGAAGAGCCCNCTCTGTGCAATTCCG 462
Db 61 GGCTCCAGATGGAGGTGGATGTCTGTGCCATAAAGAGGCCGCTGGCAGTCCCTTC 120
QY 463 AGCTACACAAAGAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCA 522
Db 121 CGAGCTGTCTAGGAGAGAGCATGGCCGTGCGAGAGGCCACATGCCCTGGATTGCA 180
QY 523 TTTCGACGACAGACCTTCACACATTCAGAAACCTTAATAAACATGTCTTAATGCAAC 582
Db 181 TTTCGACGACAGTCTTCTCTCAGCGGAGGATCTCAGTCAGCACGTGCTGTCGACAC 240
QY 583 CGGCTACCTCTGTGAACACGACGATCTTCTGGGTGAAGCAGAGTATCTCAGTCCGCTT 642
Db 241 CGGCCACCTCTCGAGCCAGCTGTCTCTGTGTGGAGGCCGAGTACTTAAGTCCCTTT 300
QY 643 GATAAAGTCAAGTCGCAACAGAACTCCCAAGGAAAAAGAAATTGCAAGGA---AAATGAA 699
Db 301 GATAAAGCTCTGGAGCCAAACAGAGCCAGCATTTGGAGAGAGTGGCGAAGACCCCGAG 360
QY 700 TTTAGCTGTGAGGTATGTGGGACAGACATTTAGATGCGCTTTTGTGATGTTGAGATC 759
Db 361 TTGAGCTGTGATGTGTGTGGGACAGACATTCACAGTGGCTTTTGTGATGTTGAGAG 420
QY 760 AGAACACACAAGATTTCTTCACTTACGGGTGTACATGTGCGGAAGAGATTCAGAGAG 819
Db 421 AAGAGCATAAAGGACTCTCTTACGTATGGGTGACAGATGTGCGGGAGGAGATTCAGG 480
QY 820 CTTTGGTTTCTTAAATAATCACATGCGGACRCATATATGCAAAATCGGGGCCAGAA 879
Db 481 CGTGTCTCTGAGAGACCAACATGCGGACACACATGCGCAAGTCTGGCACCAGAGCAAG 540
QY 880 CTGACGAAGGCTTGGAGAGTAGTTCAGCAAGGATCAACAGGTCGTTCAGGTGACCGG 939
Db 541 CTTCAGCAAGGCATGGAG---AGTCCAGTCAACATCAATGAAGTGGTCCAGCGCACGCC 597
QY 940 GCCGAGCATCTCTCTCTTGCATAAATCTGCATGTTTGTGGCTTCTTATTTCCAAAT 999
Db 598 CTGGAGCATCTCCACGCCCTTACAGATCTGATGTTCTGCGGCTTCTCTTCCAAAT 657
QY 1000 AAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAATCTGTTTCCGTACC 1059
Db 658 AAGCAGAGCTCATTTGAGCAGACCAAGGTTCAAGCCCAAGAAACTGTCCCCAGTCCAGC 717
QY 1060 AGCGGCGAGACAGACTCTCCACAGAGGAGAAATGCGTCTCTGAGGGAGAGATTCCT 1119
Db 718 AACGTTGCCCTGTGATGACCAACCGAGAGGAACCCACGTCCCGAGGGAAGAGTTGCT 777
QY 1120 TTGTTCAACTTGGACACCAAAATCTCACCTGAAACGGGGGAAGACCTGTCCAGATGCAT 1179
Db 778 TTTTGAATCTGAGACCCAGATCAACTGCAGTAGTACAGTGAAGCCCATGACCTGCA 837
QY 1180 CTTCAAGTCTGATCCGTTTCCACACCTTCCAGGCTTGGCAKCTGGCTACCAAGAAAG 1239
Db 838 CTTCAAGTCTGAGCCGTTTCCACCTTACCAGCATGCGAGTTGGCTTACCAAGAAAG 897
QY 1240 GCCATTTCGCC---AAGAAAGTGAAGAAATTTGGGGCAAGAGGGAGACCGACACAG 1296
Db 898 GCCGTTGGCCAGGAGAGGTGAAGAGTCAAGAGTCAAGAGGAAGGAGCAGACATGAC 957
QY 1297 TCGAGTTCGAGAGAGGAGCTTGGAG----- 1321
Db 958 TCATGCTCAGAGAAAGAGGAACTAGGGGAAATATGGGTGGGGGTAAGCGGAAAGGTTCT 1017
QY 1322 -----AAACAAATAAGAAACCATTTGTGAGGCCCTCTCGCAAGAGAAAGAGAGTGC 1371

```
Db      1018 GGAAAGTCCAAACAAGTAAAGCAGTTGTCCAGGTCTCTCCCAAGACAAAGAGAGCCT 1077
QY      1372 AAACACTCCACGCGGAGCGCCTCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAG 1431
Db      1078 AGACATGCTAATAGTAGTGAAGTGCCCTTCTGGGGATAGTGACCCCAAGTTGTCCAGTAGCAAG 1137
QY      1432 GAGAAGCCCACTCACTGCTCCGAGTGGGCAAGCTTTCAGAACCTTACCACCACTGGTGC 1491
Db      1138 GAGAAGCCCACTCACTGCTGTTCTGAGTGCAGCAAGCCCTTCAGGACATACCACCACTCGTC 1197
QY      1492 TTGCACTCCAGGGTCC 1507
Db      1198 CTGCACTCGAGGGTGC 1213

RESULT 13
AAK81095
ID AAK81095 standard; DNA; 566 BP.
AC AAK81095;
XX
XX
XX 07-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35907.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX
XX WO200157182-A2.
PN
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226688P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
PR      01-SEP-2000; 2000US-0229343P.
PR      01-SEP-2000; 2000US-0229344P.
PR      01-SEP-2000; 2000US-0229345P.
PR      05-SEP-2000; 2000US-0229509P.
PR      06-SEP-2000; 2000US-0229513P.
PR      06-SEP-2000; 2000US-0230437P.
PR      06-SEP-2000; 2000US-0230438P.
PR      08-SEP-2000; 2000US-0231242P.
PR      08-SEP-2000; 2000US-0231243P.
PR      08-SEP-2000; 2000US-0231244P.
PR      08-SEP-2000; 2000US-0231413P.
PR      08-SEP-2000; 2000US-0231414P.
PR      08-SEP-2000; 2000US-0232080P.
PR      08-SEP-2000; 2000US-0232081P.
PR      12-SEP-2000; 2000US-0231968P.
PR      14-SEP-2000; 2000US-0232397P.
PR      14-SEP-2000; 2000US-0232398P.
PR      14-SEP-2000; 2000US-0232399P.
PR      14-SEP-2000; 2000US-0232400P.
PR      14-SEP-2000; 2000US-0232401P.
PR      14-SEP-2000; 2000US-0233063P.
PR      14-SEP-2000; 2000US-0233064P.
PR      14-SEP-2000; 2000US-0233065P.
PR      21-SEP-2000; 2000US-0234223P.
PR      21-SEP-2000; 2000US-0234274P.
PR      25-SEP-2000; 2000US-0234997P.
PR      25-SEP-2000; 2000US-0234998P.
PR      26-SEP-2000; 2000US-0235484P.
PR      27-SEP-2000; 2000US-0235834P.
PR      27-SEP-2000; 2000US-0235836P.
PR      29-SEP-2000; 2000US-0236327P.
PR      29-SEP-2000; 2000US-0236367P.
PR      29-SEP-2000; 2000US-0236368P.
PR      29-SEP-2000; 2000US-0236369P.
PR      29-SEP-2000; 2000US-0236370P.
PR      02-OCT-2000; 2000US-0236802P.
PR      02-OCT-2000; 2000US-0237037P.
PR      02-OCT-2000; 2000US-0237038P.
PR      02-OCT-2000; 2000US-0237039P.
PR      13-OCT-2000; 2000US-0237040P.
PR      13-OCT-2000; 2000US-0239335P.
PR      13-OCT-2000; 2000US-0239337P.
PR      20-OCT-2000; 2000US-0240960P.
PR      20-OCT-2000; 2000US-0241221P.
PR      20-OCT-2000; 2000US-0241785P.
PR      20-OCT-2000; 2000US-0241786P.
PR      20-OCT-2000; 2000US-0241787P.
PR      20-OCT-2000; 2000US-0241808P.
PR      20-OCT-2000; 2000US-0241809P.
PR      20-OCT-2000; 2000US-0241826P.
PR      01-NOV-2000; 2000US-0244617P.
PR      08-NOV-2000; 2000US-0246474P.
PR      08-NOV-2000; 2000US-0246475P.
PR      08-NOV-2000; 2000US-0246476P.
PR      08-NOV-2000; 2000US-0246477P.
PR      08-NOV-2000; 2000US-0246478P.
PR      08-NOV-2000; 2000US-0246523P.
PR      08-NOV-2000; 2000US-0246524P.
PR      08-NOV-2000; 2000US-0246525P.
PR      08-NOV-2000; 2000US-0246526P.
PR      08-NOV-2000; 2000US-0246527P.
PR      08-NOV-2000; 2000US-0246528P.
PR      08-NOV-2000; 2000US-0246529P.
PR      08-NOV-2000; 2000US-0246609P.
PR      08-NOV-2000; 2000US-0246610P.
PR      08-NOV-2000; 2000US-0246611P.
PR      08-NOV-2000; 2000US-0246613P.
PR      17-NOV-2000; 2000US-0249207P.
PR      17-NOV-2000; 2000US-0249208P.
PR      17-NOV-2000; 2000US-0249209P.
PR      17-NOV-2000; 2000US-0249210P.
PR      17-NOV-2000; 2000US-0249211P.
PR      17-NOV-2000; 2000US-0249212P.
```

PR 17-NOV-2000; 2000US-02492113P.
PR 17-NOV-2000; 2000US-02492114P.
PR 17-NOV-2000; 2000US-02492115P.
PR 17-NOV-2000; 2000US-02492116P.
PR 17-NOV-2000; 2000US-02492117P.
PR 17-NOV-2000; 2000US-02492118P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 35907; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM821170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (II)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 566 BP; 170 A; 127 C; 134 G; 135 T; 0 U; 0 Other;

Query Match 34.7%; Score 523.2; DB 4; Length 566;
Best Local Similarity 96.1%; Pred. No. 2.8e-159;
Matches 545; Conservative 1; Mismatches 20; Indels 1; Gaps 1;

QY 430 TGTCAATGAAGGGCCGCTGTGTTCATTCGAGCTACACAAAGAAAAATGTCATC 489
DB 1 TGTCAATGAAGGGACCGCTGTGTTCATTCGAGCTACACAAAGAAAAATGTC-ATC 59
QY 490 CGAATTCGAGGGGAATATGCCCTTGATTCGATTCGAGCCAGACCTTCACACATTC 549
DB 60 CAATTCGAGGGGTATATGCCCTTGATTCGATTCGAGCCAGACCTTCACATTC 119
QY 550 GAAGACCTTAATAACATGTCTTAATGCAACACCGGCCCTACCCCTGTGTGAACGACGTT 609
DB 120 GAAGACCTTAATAACATGTCTTAATGCAACACCGGCCCTACCCCTGTGTGAACGACGTT 179

QY 610 CTTGGGTTGAAGCAGAGATATCTCAGTCGGCTTGATATAAAGTCAAGTCGGACAGACCT 669
DB 180 CTTGGGTTGAAGCAGAGATATCTCAGTCGGCTTGATATAAAGTCAAGTCGGACAGACCT 239
QY 670 CCCAAGGAAAAGATTGCAAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCGAGACATTT 729
DB 240 CCCAAGGAAAAGATTGCAAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCGAGACATTT 299
QY 730 AGAGTCGCTTTTGTGATTTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGG 789
DB 300 AGAGTCGCTTTTGTGATTTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGG 359
QY 790 TGTAAACATGTGGGAGAAAGATTCAAGGAGCTTGGTTTCTTAAAAATCACATGCGGACR 849
DB 360 TGTAAACATGTGGGAGAAAGATTCAAGGAGCTTGGTTTCTTAAAAATCACATGCGGACA 419
QY 850 CATTAATGCAAAATCGGGGGCCAGAAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCA 909
DB 420 CATTAATGCAAAATCGGGGGCCAGAAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCA 479
QY 910 ACGATCAACGAGGTCGTCCAGGTGCACGCGCCGAGAGACATCTCTCTCTTGGCAAAATC 969
DB 480 ACGATCAACGAGGTCGTCCAGGTGCACGCGCCGAGAGACATCTCTCTCTTGGCAAAATC 539
QY 970 TGCATGGTTTGTGGCTTCTTATTTCCA 996
DB 540 TGCATGGTTTGTGGCTTCTTATTTCCA 566

RESULT 14
AAK81093
ID AAK81093 standard; DNA; 14906 BP.
AC AAK81093;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35905.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US0011354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 26-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 18:22:52 ; Search time 171.067 Seconds
(without alignments)
14414.621 Million cell updates/sec

Title: US-08-731-499-3
Perfect score: 1507
Sequence: 1 GCAGGTTCGTGGATTGACT.....GGCTTGCACATCCAGGGTCC 1507

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/PTUS COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1504.8	99.9	1507	2	US-08-680-395-3
2	1504.8	99.9	1507	4	US-08-892-695-3
3	1387.2	92.1	10365	4	US-08-892-695-9
4	1387.2	92.1	20022	4	US-09-949-016-12604
5	1387.2	92.1	20023	4	US-09-949-016-16004
6	1337.2	88.7	5632	3	US-09-560-594-3
7	1337.2	88.7	5632	4	US-09-949-016-862
8	1337.2	88.7	5632	4	US-09-949-016-4262
9	1108.6	73.6	3186	4	US-08-892-695-10
10	550.8	36.5	601	4	US-09-949-016-34699
11	550.8	36.5	601	4	US-09-949-016-151879
12	454.8	30.2	601	4	US-09-949-016-34700
13	454.8	30.2	601	4	US-09-949-016-151880
14	197.2	13.1	298	4	US-09-513-999C-11828
15	68.8	4.6	2765	4	US-09-620-312D-61
16	68.8	4.6	2799	4	US-09-949-016-4961
17	68.8	4.6	3039	4	US-09-949-016-1038
18	68.8	4.6	12323	4	US-09-949-016-16703
19	50.2	3.3	7218	1	US-08-232-463-14
20	48.6	3.2	601	4	US-09-949-016-27493
21	48.6	3.2	601	4	US-09-949-016-92521
22	48.6	3.2	1460	4	US-09-774-528-413
23	48.6	3.2	1820	4	US-09-949-016-2645
24	48.6	3.2	1839	4	US-09-949-016-493
25	48.6	3.2	23174	4	US-09-949-016-14387
26	48.6	3.2	23187	4	US-09-949-016-12235
27	48	3.2	2378	4	US-09-949-016-3823

28	48	3.2	15632	4	US-09-949-016-15565	Sequence 15565, A
29	47.2	3.1	1563	4	US-09-016-434-1386	Sequence 1386, Ap
30	47	3.1	2752	4	US-09-949-016-3563	Sequence 3563, Ap
31	47	3.1	18824	4	US-09-949-016-15305	Sequence 15305, A
32	46.8	3.1	831	4	US-09-016-434-725	Sequence 725, App
33	46.4	3.1	656	4	US-09-016-434-1406	Sequence 1406, Ap
34	46.2	3.1	857	4	US-09-016-434-892	Sequence 892, App
35	44.2	2.9	605	4	US-09-621-976-2809	Sequence 2809, Ap
36	44	2.9	293	4	US-09-016-434-23	Sequence 23, Appl
37	43.4	2.9	2402	4	US-09-949-016-1710	Sequence 1710, Ap
38	43.4	2.9	2407	4	US-09-949-016-489	Sequence 489, App
39	43.4	2.9	30867	4	US-09-949-016-12231	Sequence 12231, A
40	43.4	2.9	30868	4	US-09-949-016-13452	Sequence 13452, A
41	43.2	2.9	2555	4	US-09-620-312D-1050	Sequence 1050, Ap
42	43	2.9	241	4	US-09-016-434-746	Sequence 746, App
43	43	2.9	2920	4	US-09-620-312D-1084	Sequence 1084, Ap
44	41.4	2.7	1988	4	US-09-620-312D-1080	Sequence 1080, Ap
45	40.6	2.7	1189	4	US-09-016-434-1400	Sequence 1400, Ap

ALIGNMENTS

RESULT 1
US-08-680-395-3
; Sequence 3, Application US/08680395
; Patent No. 5892010
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W.
; APPLICANT: Collins, Colin
; APPLICANT: Hwang, Soo-in
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowbel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: Genes from the 20q13 Amplicon and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,395
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-06890005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1507 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..1507
; OTHER INFORMATION: /note= "cDNA clone cc49 of 6-7kb
; OTHER INFORMATION: transcript with homology to C2H2 zinc
; OTHER INFORMATION: finger genes"

US-08-680-395-3

Query Match 99.9%; Score 1504.8; DB 2; Length 1507;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 0; Indels 0;

Qy	1	GCAGGTTGCTGGGATTGACCTTCTTGCTCAATTGAAACACTCATTTCAATGAGAGCAAAAGAG	60
Db	1	GCAGGTTGCTGGGATTGACCTTCTTGCTCAATTGAAACACTCATTTCAATGAGAGCAAAAGAG	60
Qy	61	CACATAATGCTTTGTGCTGATTCAATATTGAATCGAGGCAATGGGAACCCCTGTATGCCTTG	120
Db	61	CACATAATGCTTTGTGCTGATTCAATATTGAATCGAGGCAATGGGAACCCCTGTATGCCTTG	120
Qy	121	TTTGTGAAAAGAACCAAGTGCAGACCATCATCTGAGCTTCTTAAAAGTTTCGAAGAAGTTAGAG	180
Db	121	TTTGTGAAAAGAACCAAGTGCAGACCATCATCTGAGCTTCTTAAAAGTTTCGAAGAAGTTAGAG	180
Qy	181	GACTATACACTTTCTTTTGAACCTTTTATAATAATATTGCTCTGGTTTGGTAAACCCAGG	240
Db	181	GACTATACACTTTCTTTTGAACCTTTTATAATAATATTGCTCTGGTTTGGTAAACCCAGG	240
Qy	241	ACTGTTAGAGGTTGAGTGACAGGTTCTTACAGTGGCCCTTAATTCAACTCCAGAAATTTGCC	300
Db	241	ACTGTTAGAGGTTGAGTGACAGGTTCTTACAGTGGCCCTTAATTCAACTCCAGAAATTTGCC	300
Qy	301	AACGGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAAATGCGCAACTCAATCCCT	360
Db	301	AACGGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAAATGCGCAACTCAATCCCT	360
Qy	361	CTTAATGTCATGGATGGCCAAAGTAGATTGGCAGCTCTCTTGCAGTCCGATGGAGATG	420
Db	361	CTTAATGTCATGGATGGCCAAAGTAGATTGGCAGCTCTCTTGCAGTCCGATGGAGATG	420
Qy	421	GAGATGCTTTGTCAAATGAAAGGGCCNCTGTTGTCAAATCCGAGCTACACAAAGAAAAA	480
Db	421	GAGATGCTTTGTCAAATGAAAGGGCCNCTGTTGTCAAATCCGAGCTACACAAAGAAAAA	480
Qy	481	ATGTCAATCCGAATTCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCAGCCAGACTTC	540
Db	481	ATGTCAATCCGAATTCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCAGCCAGACTTC	540
Qy	541	ACACATTCAGAGACCTTAATAACATGTCCTTAATGCAACACCGGCTCACTCTGTGAA	600
Db	541	ACACATTCAGAGACCTTAATAACATGTCCTTAATGCAACACCGGCTCACTCTGTGAA	600
Qy	601	CCAGCAGTTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGTTGATAAAAGTCAAGTCCGA	660
Db	601	CCAGCAGTTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGTTGATAAAAGTCAAGTCCGA	660
Qy	661	ACAGAACCTCCCAAGGAAAAGATTGCAAGGAAAAATGAATTTAGCTGTGAGGTATGTGGG	720
Db	661	ACAGAACCTCCCAAGGAAAAGATTGCAAGGAAAATGAATTTAGCTGTGAGGTATGTGGG	720
Qy	721	CAGACATTTAGAGTCGCTTTTGTGATTGAGATCCACATGAGAACACACAAGATTTCTTC	780
Db	721	CAGACATTTAGAGTCGCTTTTGTGATTGAGATCCACATGAGAACACACAAGATTTCTTC	780
Qy	781	ACTTACGGGTGTAACTGTCGGNAGAGATTCNAGGAGCCTTGTTTCTTAAATATCAC	840
Db	781	ACTTACGGGTGTAACTGTCGGNAGAGATTCNAGGAGCCTTGTTTCTTAAATATCAC	840
Qy	841	ATCGGACRCATATGCGCAAATCGGGGGCCAGAGCAAACTGCAGCAAGGCTTTGGAGAGT	900
Db	841	ATCGGACRCATATATGCGCAAATCGGGGGCCAGAGCAAACTGCAGCAAGGCTTTGGAGAGT	900
Qy	901	AGTCCAGCAGGATCAACGAGGTGTCAGGTGACCGGGCCGAGAGCATCTCTCTCCT	960
Db	901	AGTCCAGCAACGATCAACGAGGTGTCAGGTGACCGGGCCGAGAGCATCTCTCTCCT	960
Qy	961	TGCAAAATCTGCATGGTTTGTGGCTTCTCTATTTTCCAAATAAGAGAACTCTAATTGAGCAC	1020
Db	961	TGCAAAATCTGCATGGTTTGTGGCTTCTCTATTTTCCAAATAAGAGAACTCTAATTGAGCAC	1020

RESULT 2

```

US-08-892-695-3
; Sequence 3, Application US/08892695A
; Patent No. 6808678
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20013 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892.695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cc49
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (447)
; OTHER INFORMATION: N is A, G, T, or U
US-08-892-695-3

```

Query Match		99.9%;	Score 1504.8;	DB 4;	Length 1507;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1507;		Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	GCAGGTTGCTGGGATGACTCTTCTGCTCAATGAAACACTCATTCATGAGACAAAGAG	60		
DB	1	GCAGGTTGCTGGGATGACTCTTCTGCTCAATGAAACACTCATTCATGAGACAAAGAG	60		
QY	61	CACAAATGCTTGTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCCCTGTATGCTTGT	120		
DB	61	CACAAATGCTTGTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCCCTGTATGCTTGT	120		
QY	121	TTTGTGGAAGAACACAGTACACATCTGAGCTTCTTAAAGTTTGAAGTTAGAG	180		
DB	121	TTTGTGGAAGAACACAGTACACATCTGAGCTTCTTAAAGTTTGAAGTTAGAG	180		
QY	181	GACTATACACTTCTTTTGAATTTTATATTAATTAATTTGCTGCTGTTTGAACCCAGG	240		
DB	181	GACTATACACTTCTTTTGAATTTTATATTAATTAATTTGCTGCTGTTTGAACCCAGG	240		
QY	241	ACTGTTAGAGGCTGAGTGACAGGTCTTACAGTGGCCCTTAATCCAACTCCAGAAATTTGCC	300		
DB	241	ACTGTTAGAGGCTGAGTGACAGGTCTTACAGTGGCCCTTAATCCAACTCCAGAAATTTGCC	300		
QY	301	AACGGAACCTTTGAGATTTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCT	360		
DB	301	AACGGAACCTTTGAGATTTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCT	360		
QY	361	CTTAATGATACAGGATGGCAAGAGTGATGACAGCTCTCTTGGCCAGTCCGATGAGATG	420		
DB	361	CTTAATGATACAGGATGGCAAGAGTGATGACAGCTCTCTTGGCCAGTCCGATGAGATG	420		
QY	421	GAGATGCCCTTGTCAATGAAAGGCCCNCTGTTGTCAATTCGAGCTTACACAAAGAAAAA	480		
DB	421	GAGATGCCCTTGTCAATGAAAGGCCCNCTGTTGTCAATTCGAGCTTACACAAAGAAAAA	480		
QY	481	ATGTCAATCCGAATCGAGGGGAATATGCTTGGATTTGATGCTTTCGAGCCAGACCTTC	540		
DB	481	ATGTCAATCCGAATCGAGGGGAATATGCTTGGATTTGATGCTTTCGAGCCAGACCTTC	540		
QY	541	ACACATTCAGAGACCTTAAATAACATGCTTAAATGCAACACCGGCTTACCTCTGTGAA	600		
DB	541	ACACATTCAGAGACCTTAAATAACATGCTTAAATGCAACACCGGCTTACCTCTGTGAA	600		
QY	601	CCAGCAGTCTTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCA	660		
DB	601	CCAGCAGTCTTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCA	660		
QY	661	ACAGAACTCCCAAGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGG	720		
DB	661	ACAGAACTCCCAAGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGG	720		
QY	721	CAGACATTTAGAGTCCGCTTGTGATGTTGATGCTTCAATGAGACACACAAAGTTCTTTC	780		
DB	721	CAGACATTTAGAGTCCGCTTGTGATGTTGATGCTTCAATGAGACACACAAAGTTCTTTC	780		
QY	781	ACTTACGGGTGTAACATGTGCGGAAGAGTTCAAGGAGCCTTGTGTTCTTAAATATCAC	840		
DB	781	ACTTACGGGTGTAACATGTGCGGAAGAGTTCAAGGAGCCTTGTGTTCTTAAATATCAC	840		
QY	841	ATGCGGACRCAATATGGCAAAATCGGGGCCAGAACCTCAGCAAGGCTTGGAGAGT	900		
DB	841	ATGCGGACRCAATATGGCAAAATCGGGGCCAGAACCTCAGCAAGGCTTGGAGAGT	900		
QY	901	AGTCAGCAACATCAACAGAGTCTGTCAGGTGTCAGGGCCGAGAGCATCTCTCTCT	960		
DB	901	AGTCAGCAACATCAACAGAGTCTGTCAGGTGTCAGGGCCGAGAGCATCTCTCTCT	960		
QY	961	TGCAAAATCTGATGTTTGGCTTCTTATTTCAAAATGAAAGAGTCTAATTCAGGAC	1020		
DB	961	TGCAAAATCTGATGTTTGGCTTCTTATTTCAAAATGAAAGAGTCTAATTCAGGAC	1020		
QY	1021	CGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCGAGCGCGCAGACACTCTCCA	1080		

DB	1021	CGCAAGGTGCACACCAAAAAAACTGCTTTCGGTACCGAGCGCGCAGACACTCTCCA	1080		
QY	1081	CAAGGAGGAATGCGGCTCTCGAGGGAGGACTTCTTGCAGTTGTTCAACTTGAGACCAAAA	1140		
DB	1081	CAAGGAGGAATGCGGCTCTCGAGGGAGGACTTCTTGCAGTTGTTCAACTTGAGACCAAAA	1140		
QY	1141	TCTCACTCCCTGAAACGGGGAGAGCCTGTGATGATGATCCCTCAGCTCGATCGTTTCC	1200		
DB	1141	TCTCACTCCCTGAAACGGGGAGAGCCTGTGATGATGATCCCTCAGCTCGATCGTTTCC	1200		
QY	1201	ACCTTCAGGCTTGGCAKCTGCTACCAAGAAAGTTGCCATTTGCCAAGAGTGAAG	1260		
DB	1201	ACCTTCAGGCTTGGCAKCTGCTACCAAGAAAGTTGCCATTTGCCAAGAGTGAAG	1260		
QY	1261	GAATTTGGGCGCAAGAGGAGCACCCGACACAGGATTCGAGTTCCGAGAGGAGCTTGA	1320		
DB	1261	GAATTTGGGCGCAAGAGGAGCACCCGACACAGGATTCGAGTTCCGAGAGGAGCTTGA	1320		
QY	1321	GAAACAAATAGAACCATTTGTGAGGCTCTCGCAAGAGAAAGAGTGCACAACTCC	1380		
DB	1321	GAAACAAATAGAACCATTTGTGAGGCTCTCGCAAGAGAAAGAGTGCACAACTCC	1380		
QY	1381	CAGGCGAAGCGCCCTCGTGAGACGGGATCCCAAGTTACCCAGTAGCAAGAGAGCCC	1440		
DB	1381	CAGGCGAAGCGCCCTCGTGAGACGGGATCCCAAGTTACCCAGTAGCAAGAGAGCCC	1440		
QY	1441	ACTCACTGCTCGAGTGCGGCAAGCTTTTCAAGACCTTACACAGCTGCTTTCGACTCC	1500		
DB	1441	ACTCACTGCTCGAGTGCGGCAAGCTTTTCAAGACCTTACACAGCTGCTTTCGACTCC	1500		
QY	1501	AGGGTCC 1507			
DB	1501	AGGGTCC 1507			

RESULT 3

US-08-892-695-9

Sequence 9, Application US/08892695A

Patent No. 6808878

GENERAL INFORMATION:

APPLICANT: Gray, Joe W

APPLICANT: Collins, Collin

APPLICANT: Hwang, Soo In

APPLICANT: Godfrey, Tony

APPLICANT: Kowel, David

APPLICANT: Rommens, Johanna

TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES

FILE REFERENCE: 2500.124US3

CURRENT APPLICATION NUMBER: US/08/892,695A

CURRENT FILING DATE: 1997-07-15

EARLIER APPLICATION NUMBER: 08/785,532

EARLIER FILING DATE: 1997-01-17

EARLIER APPLICATION NUMBER: 08/731,499

EARLIER FILING DATE: 1996-10-16

EARLIER APPLICATION NUMBER: 08/680,395

EARLIER FILING DATE: 1996-07-15

NUMBER OF SEQ ID NOS: 59

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 9

LENGTH: 10365

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Genomic

OTHER INFORMATION: Sequence encoding ZABCI

Patent No. 6808878

US-08-892-695-9

Query Match 92.1%; Score 1387.2; DB 4; Length 10365;

Best Local Similarity 97.3%; Pred. No. 0;

Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 2 CAGGTTGCGGATGACTTCTTCTCAATTGAAACACTCATTCAATGAGACAAGAGC 61
Db |||||
2424 CAGGTTGCTGGGATGACTTCTTGTCAATTGAAACACTCATTCAATGAGACAAGAGA 2483
QY 62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCCCTGTATGCTTTGT 121
Db |||||
2484 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCCCTGTATGCTTTGT 2543
QY 122 TTGTGAAAGAACGACGATGACACCATCATCTGAGCTTCTTAAAGATTGGAAGATTAGAGG 181
Db |||||
2544 TTGTGAAAGAACGACGATGACACCATCATCTGAGCTTCTTAAAGATTGGAAGATTAGAGG 2603
QY 182 ACTATACACTTTCTTTTGAACCTTTTATAATAATTTTGTCTCTGG-TTTTGGAAACCCAGG 240
Db |||||
2604 ACTATACACTTTCTTTTGAACCTTTTATAATAATAATTTTGTCTCTGGTTTITGGAAACCCAGG 2663
QY 241 ACTGTTAGA-GGGTGAGTGACAGGCTTTAC-AGTGGCCCTTAATCCAACTCCAGAAATTTGC 298
Db |||||
2664 GCTGTTAGAGGGTGAGTGACAAGTCTTACAAGTGGCCCTTATTCCACTCCAGAAATTTGC 2723
QY 299 CCAACGGAACTTTGAGATTATGCAATCGAAAGTGACAGGAAACATGCGCAACTCAATCC 358
Db |||||
2724 CCAACGGAACTTTGAGATTATGCAATCGAAAGTGACAGGAAACATGCGCAACTCAATCC 2783
QY 359 CTCCTTAATGTACATGAT-GGCCAAGAGTGATTGGCAGCTCTCTTGCCAGTCCGATGGAG 417
Db |||||
2784 CTCCTTAATGTACATGATGGGCCAGAGTGATTGGCAGCTCTCTTGCCAGTCCGATGGAG 2843
QY 418 ATGGA-GATGCCCTTGTCAATGAAAGGGCCNCTGTTGTCAATTCGAGCTACACAAAGAA 476
Db |||||
2844 ATGGAGGATGCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCGGAGCTACACAAGAA 2903
QY 477 AAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGCTTCTGAGCCAGAC 536
Db |||||
2904 AAAATGTG-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGCTTCTGAGCCAGAC 2962
QY 537 CTTACACATTCAGAAAGACTTAAATAACATGCTTAAATGCAACACGGCCCTACCCCTCTG 596
Db |||||
2963 CTTACACATTCAGAAAGACTTAAATAACATGCTTAAATGCAACACGGCCCTACCCCTCTG 3022
QY 597 TGAACCAAGCTTCTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 656
Db |||||
3023 TGAACCAAGCTTCTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 3082
QY 657 GCGAACAGAACTTCCAGAGAAAGAAATGCAAGAAATGAATTTAGCTGTGAGGTATG 716
Db |||||
3083 GCGAACAGAACTTCCAGAGAAAGAAATGCAAGAAATGAATTTAGCTGTGAGGTATG 3142
QY 717 TGGGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTC 776
Db |||||
3143 TGGGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTC 3202
QY 777 TTTCACTTACGGGTGAACATGTGCGGAAGAGATTCAGAGAGCCTTGCTTTCTTAAAAA 836
Db |||||
3203 TTTCACTTACGGGTGAACATGTGCGGAAGAGATTCAGAGAGCCTTGCTTTCTTAAAAA 3262
QY 837 TCACATGCGGACRCATATGCAATTCGGGGCCAGAGCAAACTGCAAGAGGCTTGA 896
Db |||||
3263 TCACATGCGGACACATAATGCGCAATTCGGGGCCAGAGCAAACTGCAAGAGGCTTGA 3322
QY 897 GAGTAGTCAGCAACGATCAACAGAGTCTGTCAGGTGCAACGGCCGAGAGCATCTCTCT 956
Db |||||
3323 GAGTAGTCAGCAACGATCAACAGAGTCTGTCAGGTGCAACGGCCGAGAGCATCTCTCT 3382
QY 957 TCCTTGAATAATCTGCATGTTTGTGGCTTCTATTTCCAAATAAGAAAGTCTAATTTGA 1016
Db |||||
3383 TCCTTGAATAATCTGCATGTTTGTGGCTTCTATTTCCAAATAAGAAAGTCTAATTTGA 3442
QY 1017 GCACCGGAGGTGACACCAAAACCTGCTTTTCGTTACCGAGCAGCGCGGACAGATC 1076
Db |||||
3443 GCACCGGAGGTGACACCAAAACCTGCTTTTCGTTACCGAGCAGCGCGGACAGATC 3502
QY 1077 TCCAAGGAGGAATGCGGCTCTCGAGGAGGACTTCTCTGCAAGTTGTTTCAACTTGAGACC 1136

Db 3503 TCCACAGAGGAATGCCGTCTCGAGGAGGACTTCTCTGAGTTGTTCAACTTGAGACC 3562
QY 1137 AAAATCTCACCTGAAACCGGGAAGAACCTGTTCAGATGCAATCCCTCAGCTCGATCCGTT 1196
Db |||||
3563 AAAATCTCACCTGAAACCGGGAAGAACCTGTTCAGATGCAATCCCTCAGCTCGATCCGTT 3622
QY 1197 CACCACTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGATTGCAATTTGCCAAGAGT 1256
Db |||||
3623 CACCACTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGATTGCAATTTGCCAAGAGT 3682
QY 1257 GAAGGAATTTGGGGAAGAGGAGCACCAGCAAAACGATTCGAGTTCGGAAGAGGAGCT 1316
Db |||||
3683 GAAGGAATTTGGGGAAGAGGAGCACCAGCAAAACGATTCGAGTTCGGAAGAGGAGCT 3742
QY 1317 TGGAGAAAATAAGAACCAATTTGTCAGGCTCTCGCAAGAGAAAGAGAGTGCACAA 1376
Db |||||
3743 TGGAGAAAATAAGGGCAGTTGTGCAAGGCTCTCGCAAGAGAAAGAGAGTGCACAA 3802
QY 1377 CTCACACGGCAAGCGCCCTCGTGAGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 1436
Db |||||
3803 CTCACACGGCAAGCGCCCTCGTGAGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 3862
QY 1437 GCCCACTCACTGCTCCGAGTGGCGAAAGCTTTTCAAGAACCTACCAACAGCTGCTTGA 1496
Db |||||
3863 GCCCACTCACTGCTCCGAGTGGCGAAAGCTTTTCAAGAACCTACCAACAGCTGCTTGA 3922
QY 1497 CTCAGGGTCC 1507
Db |||||
3923 CTCAGGGTCC 3933

RESULT 4

US-09-949-016-12604
; Sequence 12604, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12604
; LENGTH: 20022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12604

Query Match 92.1%; Score 1387.2; DB 4; Length 20022;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 2 CAGGTTGCTGGGATTTGACTTCTTCTCAATTGAAACACTCATTCAATGAGACAAGAGC 61
Db |||||
1951 CAGGTTGCTGGGATTTGACTTCTTCTCAATTGAAACACTCATTCAATGAGACAAGAGA 2010
QY 62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCAATTTGGAAACCCCTGTATGCTTTGT 121
Db |||||
2011 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCAATTTGGAAACCCCTGTATGCTTTGT 2070
QY 122 TTGTGGAAGAACCGAGTGACACCATCACTGAGCTTCTTAAAGATTGGAAGATTAGAGG 181
Db |||||
2071 TTGTGGAAGAACCGAGTGACACCATCACTGAGCTTCTTAAAGATTGGAAGATTAGAGG 2130

QY 182 ACTATACACCTTTCTTTTGAACCTTTTATAATAAATTTGCTCTGG-TTTTGAACCCAGG 240
DB 2131 ACTATACACCTTTCTTTTGAACCTTTTATAATAAATTTGCTCTGGTTTGTGAACCCAGG 2190
QY 241 ACTGTTAGA-GGGTAGTCAGAGTCTTAC-AGTGGCCTTAATCCAACTCCAGAAATTCG 298
DB 2191 GCTGTTAGAGGGTAGTCAGAGTCTTACAAGTGGCCCTTATTCCAACTCCAGAAATTCG 2250
QY 299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTCAGAGGAAACATGCCAACTCAATCC 358
DB 2251 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTCAGAGGAAACATGCCAACTCAATCC 2310
QY 359 CTCCTTAATGTACATGAT-GGCCAAGAGTGAATGGCAGCTCTCTGCCAGTCCGATGGAG 417
DB 2311 CTCCTTAATGTACATGATGGCCAGAGTGAATGGCAGCTCTCTGGCAGTCCGATGGAG 2370
QY 418 ATGGA-GATGCTTGTCAATGAAGGGCCNCTGTTGCAATTCGAGCTACACAAGAA 476
DB 2371 ATGGAGATGCTTGTCAATGAAGGGACCGCTGTTGTCCATTCGAGCTACACAAGAA 2430
QY 477 AAAAAATGTCATCCGAATCCAGGGGAATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 536
DB 2431 AAAAAATGTC-ATCCAAATCGAGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 2489
QY 537 CTTACACATTCAGAGACCTTAATAAACATGTCCTTAATGCAACACCGCCCTACCCCTCG 596
DB 2490 CTTACACATTCAGAGACCTTAATAAACATGTCCTTAATGCAACACCGCCCTACCCCTCG 2549
QY 597 TGAACCGAGTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 656
DB 2550 TGAACCGAGTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 2609
QY 657 GCGAACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATG 716
DB 2610 GCGAACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATG 2669
QY 717 TGGGAGACATTTAGAGTCGCTTTGATGTTGAGATCCCATGAGAACACACAAAGATTTC 776
DB 2670 TGGGAGACATTTAGAGTCGCTTTGATGTTGAGATCCCATGAGAACACACAAAGATTTC 2729
QY 777 TTTCACTTACGGGTGTAACATGTCGGAAGAGATTCAAGGAGCCTTGTTCTTAAAAA 836
DB 2730 TTTCACTTACGGGTGTAACATGTCGGAAGAGATTCAAGGAGCCTTGTTCTTAAAAA 2789
QY 837 TCACATGCGGACRCAATAATGCGAAATCGGGGCGAGAACAACTGCAGCAAGGCTTGA 896
DB 2790 TCACATGCGGACACATAATGCGAAATCGGGGCGAGAACAACTGCAGCAAGGCTTGA 2849
QY 897 GAGTAGTCAGCAACGATCAACGAGTCTGTCAGGTGTCACGGCGCGAGAGCATCTCCTC 956
DB 2850 GAGTAGTCAGCAACGATCAACGAGTCTGTCAGGTGTCACGGCGCGAGAGCATCTCCTC 2909
QY 957 TCGTTGCAAAATCTGCATGGTTTGGCTTCTTATTTCCAAATGAAGAAAGTCTAATGA 1016
DB 2910 TCGTTGCAAAATCTGCATGGTTTGGCTTCTTATTTCCAAATGAAGAAAGTCTAATGA 2969
QY 1017 GCACCGCAAGGTGCACACCAAAAAAATCGCTTTCGGTACAGCAGCGCGCAGACGACTC 1076
DB 2970 GCACCGCAAGGTGCACACCAAAAAAATCGCTTTCGGTACAGCAGCGCGCAGACGACTC 3029
QY 1077 TCCACAAGGAGGAATGCCCTCTCGAGGAGGACTTCTCGCAGTGTGTTCAACTTGAGACC 1136
DB 3030 TCCACAAGGAGGAATGCCCTCTCGAGGAGGACTTCTCGCAGTGTGTTCAACTTGAGACC 3089
QY 1137 AAAATCTCACCTGAAACCGGGAGAGGCTGTGATGTCATCCCTCAGCTCGATCCGTT 1196
DB 3090 AAAATCTCACCTGAAACCGGGAGAGGCTGTGATGTCATCCCTCAGCTCGATCCGTT 3149
QY 1197 CACCACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTGGCCATTTGCCAAGAGT 1256
DB 3150 CACCACCTTCCAGGCTTGGCAGCTGGCTACCAAGGAAGTGGCCATTTGCCAAGAGT 3209
QY 1257 GAAGGAATTTGGGCAAGAGGGAGCACCGAACCGACGATTCGAGTTCGAGAGAGGAGCT 1316

DB 3210 GAAGGAATCGGGCAAGAGGGAGCACCAACGACGATTCGAGTTCGAGAGGAGCT 3269
QY 1317 TCGAGAAACAAATAAGAACCAATTTGTGAGGCTCTTCGCAAGAGAAAGAGAGTGAACA 1376
DB 3270 TCGAGAAACAAATAAGGGCAGTTTGTGAGGCTCTTCGCAAGAGAAAGAGAGTGAACA 3329
QY 1377 CTCCACCGCGAAGCGCCCTCCGTGGAACGCGATCCCAAGTTTACCAGTAGCAAGAGAA 1436
DB 3330 CTCCACCGCGAAGCGCCCTCCGTGGAACGCGATCCCAAGTTTACCAGTAGCAAGAGAA 3389
QY 1437 GCCCACTCACTCTCTCGAGTGGGCAAGCTTTTCAGAACTTACCACCACTGCTCTTGA 1496
DB 3390 GCCCACTCACTCTCTCGAGTGGCGCAAGCTTTTCAGAACTTACCACCACTGCTCTTGA 3449
QY 1497 CTCCAGGGTCC 1507
DB 3450 CTCCAGGGTCC 3460

RESULT 5
US-09-949-016-16004
; Sequence 16004, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE OF INVENTION: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16004
; LENGTH: 20023
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16004

Query Match 92.1%; Score 1387.2; DB 4; Length 20023;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
QY 2 CAGGTTGCTGGGATTGACTTCTTGTCTCAATTGAAACACTCATTCAATGAGACAAAGAG 61
DB 1951 CAGGTTGCTGGGATTGACTTCTTGTCTCAATTGAAACACTCATTCAATGAGACAAAGAG 2010
QY 62 ACTAATGCTTTGCTGATTCATATTGAAATCGAGGCAATGGGAACCTGTATGCTTGT 121
DB 2011 ACTAATGCTTTGCTGATTCATATTGAAATCGAGGCAATGGGAACCTGTATGCTTGT 2070
QY 122 TTGTGGAAGAACACAGTGCACCATCAGCTTCTTAAATGTTTGAAGTTTGAAGAGTTAGAG 181
DB 2071 TTGTGGAAGAACACAGTGCACCATCAGCTTCTTAAATGTTTGAAGTTTGAAGAGTTAGAG 2130
QY 182 ACTATACACTTCTTTTGAACCTTTTATAATAAATTTGCTCTGG-TTTTGAACCCAGG 240
DB 2131 ACTATACACTTCTTTTGAACCTTTTATAATAAATTTGCTCTGGTTTGTGAACCCAGG 2190
QY 241 ACTGTTAGA-GGGTAGTCAGAGTCTTAC-AGTGGCCTTAATCCAACTCCAGAAATTCG 298
DB 2191 GCTGTTAGAGGGTAGTCAGAGTCTTACAAGTGGCCCTTATTCCAACTCCAGAAATTCG 2250
QY 299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTCAGAGGAAACATGCCAACTCAATCC 358
DB 2251 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTCAGAGGAAACATGCCAACTCAATCC 2310

QY 359 CTCCTTAATGTCATCGAT - GGCCAAAGAGTGATTGGCAGCTCTCTTGCAGTCCGATGGAG 417
 DB |||||
 QY 2311 CTCCTTAATGTCATCGATGGCCAGAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAG 2370
 DB |||||
 QY 418 ATGGA - GATGCCCTTGTCAATGAAAGGGCCNCTGTGTCAATTCGCGAGCTACACAAAGAA 476
 DB |||||
 QY 2371 ATGGAGGATGCTTGTCAATGAAAGGGACCGCTGTGTTCATTCGAGCTACACAAGAA 2430
 DB |||||
 QY 477 AAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTCGATTCGTCGAGCCAGAC 536
 DB |||||
 QY 2431 AAAATGTGC - ATCCAAATCGAGGGGTATATGCCCTTGGATTCGATTCGTCGAGCCAGAC 2489
 DB |||||
 QY 537 CTTCACATTCAGAGACCTTAAATAACATGCTTTAATGCAACACCGSCCTACCCCTCTG 596
 DB |||||
 QY 2490 CTTTACATTCAGAGACCTTAAATAACATGCTTTAATGCAACACCGSCCTACCCCTCTG 2549
 DB |||||
 QY 597 TGAACACGACGTTCTCGGGTTGAACGAGATATCTCAGTCGCTTGTATAAAGTCAAGT 656
 DB |||||
 QY 2550 TGAACACGACGTTCTCGGGTTGAACGAGATATCTCAGTCGCTTGTATAAAGTCAAGT 2609
 DB |||||
 QY 657 GCGAACAGAACCTCCAAAGGAAAGAAATGCAAGGAAATGAAATTTAGCTGTGAGGTATG 716
 DB |||||
 QY 2610 GCGAACAGAACCTCCAAAGGAAAGAAATGCAAGGAAATGAAATTTAGCTGTGAGGTATG 2669
 DB |||||
 QY 717 TGGGACAGACATTTAGATCGCTTTTGTGTTGATGTCACATGAGAACACACAAAGATTC 776
 DB |||||
 QY 2670 TGGGACAGACATTTAGATCGCTTTTGTGTTGATGTCACATGAGAACACACAAAGATTC 2729
 DB |||||
 QY 777 TTTTCACTTACGGGTGTAA CATGTGCGGAAGAGATTCAAGGAGCCTTGGTTTCTTAAAAA 836
 DB |||||
 QY 2730 TTTTCACTTACGGGTGTAA CATGTGCGGAAGAGATTCAAGGAGCCTTGGTTTCTTAAAAA 2789
 DB |||||
 QY 837 TCACATGCGGACRATATGCAATTCGGGGCCAGAGCAAACTGCAAGCAAGGCTTGA 896
 DB |||||
 QY 2790 TCACATGCGGACRATATGCAATTCGGGGCCAGAGCAAACTGCAAGCAAGGCTTGA 2849
 DB |||||
 QY 897 GAGTAGTCAGCAACGATCAACAGAGTCGTCCAGGTGCAACGGGGCCAGAGCATCTCTC 956
 DB |||||
 QY 2850 GAGTAGTCAGCAACGATCAACAGAGTCGTCCAGGTGCAACGGGGCCAGAGCATCTCTC 2909
 DB |||||
 QY 957 TCCTTGCAAAATCTGCATGGTTGTGGCTTCTATTTCCAAATAAAGAAAGTCTAATTGA 1016
 DB |||||
 QY 2910 TCCTTGCAAAATCTGCATGGTTGTGGCTTCTATTTCCAAATAAAGAAAGTCTAATTGA 2969
 DB |||||
 QY 1017 GCACCGCAAGGTGCACACCAAAAACCTGCTTCGGTACCAAGCAGCGCGCAGACAGATC 1076
 DB |||||
 QY 2970 GCACCGCAAGGTGCACACCAAAAACCTGCTTCGGTACCAAGCAGCGCGCAGACAGATC 3029
 DB |||||
 QY 1077 TCCCAAGGAGGAATGCCGTCTCGAGGAGGACCTTCTCGAGATTGTTCAACTTGAGACC 1136
 DB |||||
 QY 3030 TCCCAAGGAGGAATGCCGTCTCGAGGAGGACCTTCTCGAGATTGTTCAACTTGAGACC 3089
 DB |||||
 QY 1137 AAAATCTCACCTGAAACGGGGAAGAGCCTGTAGATGCATTCCTCAGCTCGATCCGTT 1196
 DB |||||
 QY 3090 AAAATCTCACCTGAAACGGGGAAGAGCCTGTAGATGCATTCCTCAGCTCGATCCGTT 3149
 DB |||||
 QY 1197 CACCACTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGAGTTCGCAATTCGCAAGAGT 1256
 DB |||||
 QY 3150 CACCACTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGAGTTCGCAATTCGCAAGAGT 3209
 DB |||||
 QY 1257 GAAGGAATTTGGGCAAGAGGAGCACCGCAACGAGATTTCGAGTTCGAGAGAGGAGCT 1316
 DB |||||
 QY 3210 GAAGGAATTTGGGCAAGAGGAGCACCGCAACGAGATTTCGAGTTCGAGAGAGGAGCT 3269
 DB |||||
 QY 1317 TGGAGAAA CAAATAAGAA CCAATTGTGAGGCTCTCGCAAGAGAAAGAGAGTGCAGAA 1376
 DB |||||
 QY 3270 TGGAGAAA CAAATAAGGGCAGTTGTGACGGCTCTCGCAAGAGAAAGAGAGTGCAGAA 3329
 DB |||||
 QY 1377 CTCACCGGCAAGCGCCTCGTGGACCGGATCCCAAGTTACCAGTACGCAAGGAGAA 1436
 DB |||||
 QY 3330 CTCACCGGCAAGCGCCTCGTGGACCGGATCCCAAGTTACCAGTACGCAAGGAGAA 3389
 DB |||||
 QY 1437 GCCCCACTCACTGCTCCGAGTGGCGAAAGCTTTTCAGAAACCTACCAAGGAGGAAATGAAATTTAGCT 1496
 DB |||||

DB 3390 GCCCACTCACTGCTCCGAGTGGCGAAAGCTTTGAGAACCTACCAACAGCTGGTCTTGCA 3449
 QY 1497 CTCAGGGTCC 1507
 DB 3450 CTCAGGGTCC 3460
 RESULT 6
 US-09-560-594-3
 ; Sequence 3, Application US/09560594
 ; Patent No. 6242590
 ; GENERAL INFORMATION:
 ; APPLICANT: Lex M. Cowser
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF ZINC FINGER PROTEIN-217 EXPRESSION
 ; FILE REFERENCE: RTS-0144
 ; CURRENT APPLICATION NUMBER: US/09/560,594
 ; CURRENT FILING DATE: 2000-04-28
 ; NUMBER OF SEQ ID NOS: 89
 ; SEQ ID NO 3
 ; LENGTH: 5632
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (272)...(3418)
 US-09-560-594-3
 Query Match 88.7%; Score 1337.2; DB 3; Length 5632;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
 QY 52 GACAAAGAGCACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCCTG 111
 DB 1 GACAAAGAGCACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCCTG 60
 QY 112 TATGCTTTGTTGTGGAAGAACAGTGCACACCATCACTGAGCTTCCTATAAAGTTTGAAG 171
 DB 61 TATGCTTTGTTGTGGAAGAACAGTGCACACCATCACTGAGCTTCCTATAAAGTTTGAAG 120
 QY 172 AAGTTAGAGGACTATACACTTTTCTTTTGAATTTTAAATTAATTTGCTCTGG -TTTT 230
 DB 121 AAGTTAGAGGACTATACACTTTTCTTTTGAATTTTAAATTAATTTGCTCTGGTTT 180
 QY 231 GGAACCCAGGACTGTTAGA -GGTGAGTGACAGGCTTTAC -AGTGGCCTTAATCCAACTC 288
 DB 181 GGAACCCAGGCTGTTAGAGGGGTGAGTGACAAGTCTTTACAAGTGGCCTTTATCCAACTC 240
 QY 289 CAGAAATGCCCAACCGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348
 DB 241 CAGAAATGCCCAACCGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
 QY 349 AACTCAATCCCTTTAATGTATCATGGAT -GGCAGAGATGATGGCAGCTCTTTGGCAG 407
 DB 301 AACTCAATCCCTTTAATGTATCATGGATGGGCGCAGAGTGTGGCAGCTCTCTTTGGCAG 360
 QY 408 TCCGATGAGATGGA -GATGCTTGTCAATGAAAGGGCCNCTGTGTTCATTTCCAGCT 466
 DB 361 TCCGATGAGATGAGGATGCTTGTCAATGAAAGGGACCCGCTGTGTTCATTTCCAGCT 420
 QY 467 ACACAAAGAAAAAATGTCATTCGAAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526
 DB 421 TACACAGAAAAAATGTC -ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT 479
 QY 527 GCAGCCAGACCTTCA CAAATTCAGAAAGCCTTTAATAAACA TGTCTTTAATGCAACCCGCG 586
 DB 480 GCAGCCAGACCTTCA CAAATTCAGAAAGCCTTTAATAAACA TGTCTTTAATGCAACCCGCG 539
 QY 587 CTACCTCTGTGGAACCAAGCAGTCTTCCGGTTTGAAGCAGAGTATCTCAGTCCGCTTGATA 646
 DB 540 CTACCTCTGTGGAACCAAGCAGTCTTCCGGTTTGAAGCAGAGTATCTCAGTCCGCTTGATA 599
 QY 647 AAAGTCAAGTGCAGAACAGAACTCCCAAGGAAAAAGAAATTCGAAGGAAAAATGAATTTAGCT 706
 DB |||||

Db 600 AAAGTCAAGTCGGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCT 659
Qy 707 GTGAGTATGTGGGACAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACAC 766
Db 660 GTGAGTATGTGGGACAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACAC 719
Qy 767 ACAAGATCTTTTCACTTACCGGTGTAACATGTGCGGAAAGAGATTCGAAGGAGCTTTGT 826
Db 720 ACAAGATCTTTTCACTTACCGGTGTAACATGTGCGGAAAGAGATTCGAAGGAGCTTTGT 779
Qy 827 TTCTTAAATACATGCGGACBACATAATGGCAATTCGGGGCCAGAACCAATTCGCAC 886
Db 780 TTCTTAAATACATGCGGACBACATAATGGCAATTCGGGGCCAGAACCAATTCGCAC 839
Qy 887 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGTCAGGTGTCAGCGGCGGAG 946
Db 840 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGTCAGGTGTCAGCGGCGGAG 899
Qy 947 GCATCTCTCTCTCTGCAAAATCTGCATGTTTGTGGCTTCTTATTTCAAAATAAGAAA 1006
Db 900 GCATCTCTCTCTCTGCAAAATCTGCATGTTTGTGGCTTCTTATTTCAAAATAAGAAA 959
Qy 1007 GTCTAATTCAGCACCGCAGAGTGTCACCAAAAAAATCTGTTTCTGATACAGCGGCGC 1066
Db 960 GTCTAATTCAGCACCGCAGAGTGTCACCAAAAAAATCTGTTTCTGATACAGCGGCGC 1019
Qy 1067 AGACAGACTCTCCACAGGAGGAAATCGCTGCTCGAGGAGGACTTCTGCGAGTTGTTCA 1126
Db 1020 AGACAGACTCTCCACAGGAGGAAATCGCTGCTCGAGGAGGACTTCTGCGAGTTGTTCA 1079
Qy 1127 ACTTGAGACCAAAATCTCACTTGAACCGGGAAGAGCTGTGATGATCATCTCCCTCAGC 1186
Db 1080 ACTTGAGACCAAAATCTCACTTGAACCGGGAAGAGCTGTGATGATCATCTCCCTCAGC 1139
Qy 1187 TCGATCCGTTACACCTTTCAGGCTTGGCACTGCTGCTACCAAGAAAGTTGCCATTT 1246
Db 1140 TCGATCCGTTACACCTTTCAGGCTTGGCACTGCTGCTACCAAGAAAGTTGCCATTT 1199
Qy 1247 GCCAAGAGTGAAGAAATTTGGGCAAGAGGAGGACCCACACGACGATTCGAGTTCCG 1306
Db 1200 GCCAAGAGTGAAGAAATTTGGGCAAGAGGAGGACCCACACGACGATTCGAGTTCCG 1259
Qy 1307 AGAAGAGCTTGGAGAAAACAAATAGAACCATTTGTGAGGCTCTCGCAGGAGAGAGA 1366
Db 1260 AGAAGAGCTTGGAGAAAACAAATAGAACCATTTGTGAGGCTCTCGCAGGAGAGAGA 1319
Qy 1367 AGTGCAACATCTCCACGGGAGGCGCTTCGTTGAGCGGATCCCAAGTTACCCAGTA 1426
Db 1320 AGTGCAACATCTCCACGGGAGGCGCTTCGTTGAGCGGATCCCAAGTTACCCAGTA 1379
Qy 1427 GCAAGGAGAGCCCACTCACTGCTCGAGTGGGCAAGCTTTCGAACTTACCACCGC 1486
Db 1380 GCAAGGAGAGCCCACTCACTGCTCGAGTGGGCAAGCTTTCGAACTTACCACCGC 1439
Qy 1487 TGGTCTTGCACTCCAGGGTCC 1507
Db 1440 TGGTCTTGCACTCCAGGGTCC 1460

RESULT 7

US-09-949-016-862
; Sequence 862, Application US/09949016
; Patent No. 6812319
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 862
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-862

Query Match 88.7%; Score 1337.2; DB 4; Length 5632;

Best Local Similarity 97.2%; Pred. No. 0;

Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

Qy 52 GACAAGAGCAGCACTAATGCTTGTGCTGATTCATATTTGAATCGAGCATTGGGAACCCCTG 111
Db 1 GACAAGAGCACTAATGCTTGTGCTGATTCATATTTGAATCGAGCATTGGGAACCCCTG 60
Qy 112 TATGCTTGTGCTGGAAGAACCAAGTGCACCATCCTGAGCTTCTTAAAGTTTGAAG 171
Db 61 TATGCTTGTGCTGGAAGAACCAAGTGCACCATCCTGAGCTTCTTAAAGTTTGAAG 120
Qy 172 AAGTTAGAGGACTATACACTTTCTTTTGAATTTTATATAAATATTTGCTCTGG-TTTT 230
Db 121 AAGTTAGAGGACTATACACTTTCTTTTGAATTTTATATAAATATTTGCTCTGGTTTT 180
Qy 231 GGAACCCAGGAGCTTTAG-GGTGAGTGACAGGCTTTAC-AGTGGCTTTAATCCAACTC 288
Db 181 GGAACCCAGGAGCTTTAGAGGGTGTGAGTGCAGAGTCTTACAAAGTGGCTTTATTTCAA 240
Qy 289 CAGAAATTCGCCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348
Db 241 CAGAAATTCGCCCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
Qy 349 AACTCAATCCCTTTTAAATGATATGAT-GGCCAAAGAGTGTGGCAGCTCTCTTCCAG 407
Db 301 AACTCAATCCCTTTTAAATGATATGATGGGCGAGAAAGTGTGGCAGCTCTCTTCCAG 360
Qy 408 TCCGATGGAGTGA- GATGCTTGTCAATGAAAGGCGCCCTGTGTCAATTTCCAGCT 466
Db 361 TCCGATGGAGTGGAGGATGCTTGTCAATGAAAGGAGCGCTGTGTTCATTTCCAGCT 420
Qy 467 ACACAAAGAAAAAATGTCAATCCGAATCGAGGGAATATGCTTGGATTTGATTTCT 526
Db 421 TACACAGAAAAAATGTC-ATCCAAATCGAGGGGTATATGCTTGGATTTGATTTCT 479
Qy 527 GCAGCCAGACTTTCACATTCAGAAAGCTTTAATAAACAATGTCTTTAATGCAACACCGGC 586
Db 480 GCAGCCAGACTTTCACATTCAGAAAGCTTTAATAAACAATGTCTTTAATGCAACACCGGC 539
Qy 587 CTACCTCTGTGAACCGAGGCTTCTCGGTTTGAAGCAGAGTATCTCAGTCCGCTTGATA 646
Db 540 CTACCTCTGTGAACCGAGGCTTCTCGGTTTGAAGCAGAGTATCTCAGTCCGCTTGATA 599
Qy 647 AAAGTCAAGTGCAGAACAGAACTCCCAAGAAAAAGAAATTCGAAGAAAAATGAATTTAGCT 706
Db 600 AAAGTCAAGTGCAGAACAGAACTCCCAAGAAAAAGAAATTCGAAGAAAAATGAATTTAGCT 659
Qy 707 GTGAGGTATGTGGGAGACATTTAGAGTCGCTTTTGAATTTGAGATTCACATGAGAACAC 766
Db 660 GTGAGGTATGTGGGAGACATTTAGAGTCGCTTTTGAATTTGAGATTCACATGAGAACAC 719
Qy 767 ACAAGATCTTTTCACTTACGGGTGTACATGTGCGGAGAGAGATTCAGAGGCTTTGCT 826
Db 720 ACAAGATCTTTTCACTTACGGGTGTACATGTGCGGAGAGAGATTCAGAGGCTTTGCT 779
Qy 827 TTTCTTAAAAATCACATGCGGACRCAATATGGCAAAATTCGGGGCCAGAACCAATTCGAGC 886
Db 780 TTTCTTAAAAATCACATGCGGACACATATATGGCAAAATTCGGGGCCAGAACCAATTCGAGC 839
Qy 887 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTTGACGCGGCGGAGA 946

```

Db      840  AAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGTGCACGCGGCCGAGA 899
Qy      947  GCATCTCTCTCTCTTGGCAAAATCTGCATGTTGGTGTGCTTCTATTTCAAATAAGAAA 1006
Db      900  GCATCTCTCTCTCTTACAAAATCTGCATGTTGGTGTGCTTCTATTTCCAAATAAGAAA 959
Qy      1007  GTCTAATTTGAGCACCAGCGAGGTGCACACCAAAAATCTGTTTCGTTACCGAGCGCGC 1066
Db      960  GTCTAATTTGAGCACCAGCGAGGTGCACCAAAAATCTGTTTCGTTACCGAGCGCGC 1019
Qy      1067  AGACAGACTCTCCCAAGGAGGAATGCGCTCTCGAGGGAGGACTTCTGCGAGTTGTTCA 1126
Db      1020  AGACAGACTCTCCCAAGGAGGAATGCGCTCTCGAGGGAGGACTTCTGCGAGTTGTTCA 1079
Qy      1127  ACTTGAGACCAAAATCTCACTCCCTGAAACGGGGAAGAGCTGTGCAGTGCAATCCCTCAGC 1186
Db      1080  ACTTGAGACCAAAATCTCACTCCCTGAAACGGGGAAGAGCTGTGCAGTGCAATCCCTCAGC 1139
Qy      1187  TCGATCCGTTTCAACACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTTGCCATTT 1246
Db      1140  TCGATCCGTTTCAACACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTTGCCATTT 1199
Qy      1247  GCCAAGAGTGAAGGAATTTGGGCAAGAGGAGCACCAGCAACACGATTCGAGTTCGG 1306
Db      1200  GCCAAGAGTGAAGGAATTCGGGCAAGAGGAGCACCAGCAACGATTCGAGTTCGG 1259
Qy      1307  AGAAGGAGCTTGGAGAAACAAATAAGAACCAATTTGTCAGGCTCTTCGCAAGAGAAAGAGA 1366
Db      1260  AGAAGGAGCTTGGAGAAACAAATAAGAGGAGTGTGTCAGGCTCTTCGCAAGAGAAAGAGA 1319
Qy      1367  AGTCAACACATCCCAACGGCAGCGCTTGGCAKCTGGCTACCAAGGAAGTTGCCATTT 1426
Db      1320  AGTCAACACATCCCAACGGCAGCGCTTGGCAKCTGGCTACCAAGGAAGTTGCCATTT 1379
Qy      1427  GCAAGGAGAGCCCACTCACTCTCCGAGTGGGCAAGGCTTTCAGAACTTACCACGTA 1486
Db      1380  GCAAGGAGAGCCCACTCACTCTCCGAGTGGGCAAGGCTTTCAGAACTTACCACGTA 1439
Qy      1487  TGGTCTTGCACTCCAGGCTCC 1507
Db      1440  TGGTCTTGCACTCCAGGCTCC 1460

```

RESULT 8

```

; Sequence 4262, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4262
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4262

```

Query Match 88.7%; Score 1337.2; DB 4; Length 5632;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

```

Qy      52  GACAAAGAGCACTAATGCTTTGCTGCTCATATTTGAATCGAGGCAATTTGGAAACCCCTG 111
Db      1  GACAAAGAGCACTAATGCTTTGCTGCTCATATTTGAATCGAGGCAATTTGGAAACCCCTG 60
Qy      112  TATGCTTTGTTTGTGGAAAGAACCAAGTGCACCACTACTGAGCTTCTTAAAGTTTCGAAG 171
Db      61  TATGCTTTGTTTGTGGAAAGAACCAAGTGCACCACTACTGAGCTTCTTAAAGTTTCGAAG 120
Qy      172  AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAAATTTGCTCTGG-TTTT 230
Db      121  AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAAATTTGCTCTGGTTTTT 180
Qy      231  GGAACCCAGGAGCTGTTAGA-GGGTGAGTGACAGGTCCTTAC-AGTGGCTTAACTCAACTC 288
Db      181  GGAACCCAGGAGCTGTTAGAAGGGTGAGTGACAGGTCCTTAAAGTGCGCTTATTTCCAATC 240
Qy      289  CAGAAATTTGCCAACCGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348
Db      241  CAGAAATTTGCCAACCGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
Qy      349  AACTCAATCCCTCTTAATGTACATGGAT-GGCAAGAGTGATTGGCAGCTCTCTTGCAG 407
Db      301  AACTCAATCCCTCTTAATGTACATGGATGGGCGCAGAGTGATTGGCAGCTCTCTTGGCAG 360
Qy      408  TCCGATGAGATGGA-GATGCTCTGTCATGAAGGGGCCNCTGTTGTCAATTCGAGCT 466
Db      361  TCCGATGAGATGAGGATGCTTGTCAATGAAGGGGCCGCTGTTGTTCCATTCGAGC 420
Qy      467  ACACAAAGAAAAAATGTCAATCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526
Db      421  TACACAAGAAAAAATGTCT-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT 479
Qy      527  GCAGCAGACCTTACACATTCAGNAGACCTTATAAACATGTCTTAATGCAACACCCGC 586
Db      480  GCAGCAGACCTTACACATTCAGNAGACCTTATAAACATGTCTTAATGCAACACCCGC 539
Qy      587  CTACCTCTGTGAACACGAGCTTCTTGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATA 646
Db      540  CTACCTCTGTGAACACGAGCTTCTTGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATA 599
Qy      647  AAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAAAGAAATTCAGGAAAAATGAATTAGCT 706
Db      600  AAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAAAGAAATTCAGGAAAAATGAATTAGCT 659
Qy      707  GTGAGTATGTGGCGAGACATTTAGAGTCCGCTTTTGTGATGTTGAGATCCACATGAACAC 766
Db      660  GTGAGTATGTGGCGAGACATTTAGAGTCCGCTTTTGTGATGTTGAGATCCACATGAACAC 719
Qy      767  ACAAGATTTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAAGATTCAAGGAGCCTTGT 826
Db      720  ACAAGATTTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAAGATTCAAGGAGCCTTGT 779
Qy      827  TTCTTAAAAATCACATGCGGACRCATATGCAAAATCGGGGGCCAGAGCAAACTGCAGC 886
Db      780  TTCTTAAAAATCACATGCGGACACATATGCAAAATCGGGGGCCAGAGCAAACTGCAGC 839
Qy      887  AAGGCTTGAGAGTAGTCCAGCAAGATCAACAGAGTGTCCAGGTGCACGCGGCGGAGA 946
Db      840  AAGGCTTGAGAGTAGTCCAGCAAGATCAACAGAGTGTCCAGGTGCACGCGGCGGAGA 899
Qy      947  GCATCTCTCTCTTTCGCAAAATCTCATGTTGTTGTTGCTTCTTATTTCCAAATAAGAAA 1006
Db      900  GCATCTCTCTCTTTCGCAAAATCTCATGTTGTTGTTGCTTCTTATTTCCAAATAAGAAA 959
Qy      1007  GTCTAATTTGAGCACCGCAAGGTGCAACCAAAAAAATCTGTTTTCGTTACCGAGCGCGC 1066
Db      960  GTCTAATTTGAGCACCGCAAGGTGCAACCAAAAAAATCTGTTTTCGTTACCGAGCGCGC 1019
Qy      1067  AGACAGACTCTCCCAAGGAGGAATGCGCTCTCGAGGGAGGACTTCTTGCAGTTGTTCA 1126
Db      1020  AGACAGACTCTCCCAAGGAGGAATGCGCTCTCGAGGGAGGACTTCTTGCAGTTGTTCA 1079
Qy      1127  ACTTGAGACCAAAATCTCACTCCCTGAAACGGGGAAGAGCTGTGCAGATGCAATCCCTCAGC 1186

```

Db 1080 ACTTGAGACCAAAATCTCAACCTGAAACGGGGAAGAGCCTGTGAGATGCATCCCTCAGC 1139
Qy 1187 TCGATCCGTTTACCACCTTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTTGGCATTT 1246
Db 1140 TCGATCCGTTTACCACCTTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTTGGCATTT 1199
Qy 1247 GCCAAGAGTGAAGGAATTTGGGCGCAAGAGGAGGACCCGACCAACGATTCGAGTTCCG 1306
Db 1200 GCCAAGAGTGAAGGAATTTGGGCGCAAGAGGAGGACCCGACCAACGATTCGAGTTCCG 1259
Qy 1307 AGAAGGAGCTTCGAGAAACAAATAGAACCATTTGTGAGGCTCTCGCAAGAGAAAGAGA 1366
Db 1260 AGAAGGAGCTTCGAGAAACAAATAGAACCATTTGTGAGGCTCTCGCAAGAGAAAGAGA 1319
Qy 1367 AGTGAACACACTCCCAACGGCGAAGCGCCCTCGGTGAGCGCGATCCCAAGTTACCCAGTA 1426
Db 1320 AGTGAACACACTCCCAACGGCGAAGCGCCCTCGGTGAGCGCGATCCCAAGTTACCCAGTA 1379
Qy 1427 GCAAGAGAGCCCACTCACTCTCGAGTGGCGCAAAAGCTTTTCAGAACTTACCACCAAGC 1486
Db 1380 GCAAGAGAGCCCACTCACTCTCGAGTGGCGCAAAAGCTTTTCAGAACTTACCACCAAGC 1439
Qy 1487 TGGTCTTGCACTCCAGGGTCC 1507
Db 1440 TGGTCTTGCACTCCAGGGTCC 1460

RESULT 9

US-08-892-695-10
; Sequence 10, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124U3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3186
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ZABCI Open
; OTHER INFORMATION: Reading Frame
US-08-892-695-10

Query Match 73.6%; Score 1108.6; DB 4; Length 3186;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1156; Conservative 3; Mismatches 28; Indels 3; Gaps 3;
Qy 320 ATGCAATCGAAGTGCAGGAAACATGCAACTCAATCCCTCTTAATGTACATGAT-GG 378
Db 1 ATGCAATCGAAGTGCAGGAAACATGCAACTCAATCCCTCTTAATGTACATGATGGG 60
Qy 379 CCAAGAGTGAATGGCAGCTCTCTTGGCAGTGGAGATGGA-GATGCCCTTGTCAATG 437
Db 61 CCAGAGTGAATGGCAGCTCTCTTGGCAGTGGAGATGGAATGGCCTTGTCAATG 120
Qy 438 AAAGGCCCNCTGTTGTCAATTCGAGCTTACAAAGAAAAAATGTCAATCCGAATCGA 497

Db 121 AAAGGACCGCTGTTGTTCCATTCCGAGCTACACAAAGAAAAAATGTC-ATCCAAATCGA 179
Qy 498 GGGGAATATGCCCTTTGGATTTGCATGTTCTGACGACGACCTTTCACACATTTCAGAAGACCT 557
Db 180 GGGGTATATGCCCTTTGGATTTGCATGTTCTGACGACGACCTTTCACACATTTCAGAAGACCT 239
Qy 558 TAATAAACATGTTTAAATGCAACACCGGCTTACCCTCTGTGAAACCAAGCAGTTCTTCGGGT 617
Db 240 TAATAAACATGTTTAAATGCAACACCGGCTTACCCTCTGTGAAACCAAGCAGTTCTTCGGGT 299
Qy 618 TGAACGACAGTATCTCAGTCCGCTTGTATATAAAGTCAAGTGCGAACGAACTCCCAAGGA 677
Db 300 TGAACGACAGTATCTCAGTCCGCTTGTATATAAAGTCAAGTGCGAACGAACTCCCAAGGA 359
Qy 678 AAAGAATTTGCAAGAAAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTCGC 737
Db 360 AAAGAATTTGCAAGAAAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTCGC 419
Qy 738 TTTTGATGTTGAGATCCACATGAGAAACACAAAGATTCTTTTCACTTACGGGTGTAACAT 797
Db 420 TTTTGATGTTGAGATCCACATGAGAAACACAAAGATTCTTTTCACTTACGGGTGTAACAT 479
Qy 798 GTGCGAAAGAGATTCAAGGAGCCTTGGTTTCTTAAAAATCAATCGCGACRCATATGG 857
Db 480 GTGCGAAAGAGATTCAAGGAGCCTTGGTTTCTTAAAAATCAATCGCGACACATAATGG 539
Qy 858 CAAATCGGGGGCCAGAACCTGACGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAA 917
Db 540 CAAATCGGGGGCCAGAACCTGACGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAA 599
Qy 918 CGAGTCTGTCCAGGTGTCACGCGCGCGAGAGCATCTCTCTCTTCAAAATCTGCATGTT 977
Db 600 CGAGTCTGTCCAGGTGTCACGCGCGCGAGAGCATCTCTCTCTTCAAAATCTGCATGTT 659
Qy 978 TTGTGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTAGGACCCGCAAGGTGCACCA 1037
Db 660 TTGTGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTAGGACCCGCAAGGTGCACCA 719
Qy 1038 AAAAACTGTTTCGGGTACAGCAGCGCGCAGACAGACTCTCCACAAAGGAGGAATGCCGTC 1097
Db 720 AAAAACTGTTTCGGGTACAGCAGCGCGCAGACAGACTCTCCACAAAGGAGGAATGCCGTC 779
Qy 1098 CTCGAGGAGGACTCTCTGCAAGTGTTCACCTTGAGACCAAAATCTCACCTTGAACCGG 1157
Db 780 CTCGAGGAGGACTCTCTGCAAGTGTTCACCTTGAGACCAAAATCTCACCTTGAACCGG 839
Qy 1158 GAAGAAGCTGTGATGATGATCCCTCAGTCCGTTCCAGTCCGTTCCAGGCTTGGA 1217
Db 840 GAAGAAGCTGTGATGATGATCCCTCAGTCCGTTCCAGTCCGTTCCAGGCTTGGA 899
Qy 1218 KCTGCTACCAAGGAAGTGGCATTTGCCAAGAGTGAAGGAATTCGGGCAAGAGG 1277
Db 900 GCTGCTACCAAGGAAGTGGCATTTGCCAAGAGTGAAGGAATTCGGGCAAGAGG 959
Qy 1278 GAGCACCGAACCAACGATTCGAGTTCGAGAGGAGCTTGGAGAAACAAATTAAGAACCA 1337
Db 960 GAGCACCGAACCAACGATTCGAGTTCGAGAGGAGCTTGGAGAAACAAATTAAGAGGAG 1019
Qy 1338 TTGTGAGGCTCTTCGAAGAGAAAGAGAGTGCAAAACATCTCCACGGGGAAGCGCCTC 1397
Db 1020 TTGTGAGGCTCTTCGAAGAGAAAGAGAGTGCAAAACATCTCCACGGGGAAGCGCCTC 1079
Qy 1398 CGTGAAGCGGATCCCAAGTTTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTG 1457
Db 1080 CGTGAAGCGGATCCCAAGTTTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTG 1139
Qy 1458 CGGCAAGCTTTCAAGAACCTACACAGCTGGTCTTTCACCTCCAGGGTCC 1507
Db 1140 CGGCAAGCTTTCAAGAACCTACACAGCTGGTCTTTCACCTCCAGGGTCC 1189

RESULT 10

```
US-09-949-016-34699/c
; Sequence 34699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34699
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34699

Query Match      36.5%; Score 550.8; DB 4; Length 601;
Best Local Similarity 96.1%; Pred. No. 8.1e-173;
Matches 572; Conservative 2; Mismatches 20; Indels 1; Gaps 1;

QY 423 GATGCTTTGTCAATGAAAGGGCCNCTGTTGTCATTCGAGCTACACAAAGAAAAAT 482
DB 594 GATGCTTTGTCAATGAAAGGGACCGCTGTTGTCATTCGAGCTACACAAAGAAAAAT 535
QY 483 GTCATTCGAAATCGAGGGGAATATGCCCTTGGAATGCAATGTTCTGCAGCCAGACCTTCAC 542
DB 534 GTC-ATCCAAATCGAGGGGTATATGCCCTTGGAATGCAATGTTCTGCAGCCAGACCTTCAC 476
QY 543 ACATTTCAGAACCTTTAATAACATGCTTAAATGCAACACCGGCCCTACCTCTGTGAACC 602
DB 475 ACATTTCAGAACCTTTAATAACATGCTTAAATGCAACACCGGCCCTACCTCTGTGAACC 416
QY 603 AGCAGTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAGTCAAGTCGGAAC 662
DB 415 AGCAGTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAGTCAAGTCGGAAC 356
QY 663 AGAACCTCCCAAGGAAAGAAATTCGAAGAAAAATGAATTTAGCTGTGAGGTATGTGGCA 722
DB 355 AGAACCTCCCAAGGAAAGAAATTCGAAGAAAAATGAATTTAGCTGTGAGGTATGTGGCA 296
QY 723 GACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAGATTTCTTTCAC 782
DB 295 GACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAGATTTCTTTCAC 236
QY 783 TTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAATCAAT 842
DB 235 TTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAATCAAT 176
QY 843 CGGACRCATATATGCAAAATCGGGGCCAGAACTGCAAGCAAGCTTCGAGCAAGCTTGGAGAGTAG 902
DB 175 CGGACACATAATGCAAAATCGGGGCCAGAACTGCAAGCAAGCTTGGAGAGTAG 116
QY 903 TCCAGCAACGATCAACAGAGGTGCTCCAGGTGCAACCGGCCGAGAGCATCTCTCTCCTTG 962
DB 115 TCCAGCAACGATCAACAGAGGTGCTCCAGGTGCAACCGGCCGAGAGCATCTCTCTCCTTA 56
QY 963 CAAAATCTGCATGTTTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAAATTGAG 1017
DB 55 CAAAATCTGCATGTTTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAAATTGAG 1

RESULT 12
US-09-949-016-34700/c
; Sequence 34700, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34700
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34700

Query Match      30.2%; Score 454.8; DB 4; Length 601;
Best Local Similarity 98.3%; Pred. No. 8.9e-141;
Matches 456; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1044 TCGTTTCGTACGAGCGCGCAGACAGACTCTCCACAGGAGGAATGCCGCTCTCGAG 1103
DB 601 TCGTTTCGTACGAGCGCGCAGACAGACTCTCCACAGGAGGAATGCCGCTCTCGAG 542
QY 1104 GGAGGACTTCTCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAA 1163
DB 541 GGAGGACTTCTCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAA 482
QY 1164 GCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCCAGCTTCCAGGCTTGGCAKCTGGC 1223
DB 481 GCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCCAGCTTCCAGGCTTGGCAKCTGGC 422
QY 1224 TACCAAGGAAGAGTTGCCATTTGCCAAGAGTGAAGAAATTTGGGCAAGAGGAGGCAC 1283
DB 421 TACCAAGGAAGAGTTGCCATTTGCCAAGAGTGAAGAAATTTGGGCAAGAGGAGGCAC 362
QY 1284 CGACACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATTAAGAACCATTTGTGC 1343
DB 361 CGACACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATTAAGAACCATTTGTGC 302
QY 1344 AGCCCTCTCGCAAGAGAAAGAGAGTGCACCACTCCCAACGCGGCAAGCGCCCTCCGTGGA 1403
DB 301 WGGGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCCAACGCGGCAAGCGCCCTCCGTGGA 242
QY 1404 CCGGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCGAGTGGCGCAA 1463
DB 241 CCGGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCGAGTGGCGCAA 182
QY 1464 AGCTTTTCAGAACTACCCAGCTGCTTTGCACTCCAGGGTCC 1507
DB 181 AGCTTTTCAGAACTACCCAGCTGCTTTGCACTCCAGGGTCC 138

RESULT 13
US-09-949-016-151880/c
; Sequence 151880, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151880
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-151880

Query Match      30.2%; Score 454.8; DB 4; Length 601;
Best Local Similarity 98.3%; Pred. No. 8.9e-141;
Matches 456; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1044 TCGTTTCGTACGAGCGCGCAGACAGACTCTCCACAGGAGGAATGCCGCTCTCGAG 1103
DB 601 TCGTTTCGTACGAGCGCGCAGACAGACTCTCCACAGGAGGAATGCCGCTCTCGAG 542
QY 1104 GGAGGACTTCTCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAA 1163
DB 541 GGAGGACTTCTCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAA 482
QY 1164 GCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCCAGCTTCCAGGCTTGGCAKCTGGC 1223
DB 481 GCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCCAGCTTCCAGGCTTGGCAKCTGGC 422
QY 1224 TACCAAGGAAGAGTTGCCATTTGCCAAGAGTGAAGAAATTTGGGCAAGAGGAGGCAC 1283
DB 421 TACCAAGGAAGAGTTGCCATTTGCCAAGAGTGAAGAAATTTGGGCAAGAGGAGGCAC 362
QY 1284 CGACACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATTAAGAACCATTTGTGC 1343
DB 361 CGACACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATTAAGAACCATTTGTGC 302
QY 1344 AGCCCTCTCGCAAGAGAAAGAGAGTGCACCACTCCCAACGCGGCAAGCGCCCTCCGTGGA 1403
DB 301 WGGGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCCAACGCGGCAAGCGCCCTCCGTGGA 242
QY 1404 CCGGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCGAGTGGCGCAA 1463
DB 241 CCGGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCGAGTGGCGCAA 182
QY 1464 AGCTTTTCAGAACTACCCAGCTGCTTTGCACTCCAGGGTCC 1507
DB 181 AGCTTTTCAGAACTACCCAGCTGCTTTGCACTCCAGGGTCC 138

RESULT 14
US-09-513-999C-11828
; Sequence 11828, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 11828
; LENGTH: 298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: 277
; OTHER INFORMATION: s=g or c
US-09-513-999C-11828

Query Match      13.1%; Score 197.2; DB 4; Length 298;
Best Local Similarity 98.5%; Pred. No. 5e-55;
```

Matches 199; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CAGGTGCTGGGATGACTTCTTCTCAATTGAACACTCATTCAATCGAGACAAGAGC 61
 Db 97 CAGGTGCTGGGATGACTTCTTGTCTCAATTGAACACTCATTCAATCGAGACAAGAGA 156

Qy 62 ACTAATGCTTTGTGCTGATTATATTTGAATCGAGGCATTGGGAACCCCTGTATGCTTGT 121
 Db 157 ACTAATGCTTTGTGCTGATTATATTTGAATCGAGGCATTGGGAACCCCTGTATGCTTGT 216

Qy 122 TTGTGGAAGAACCACTGACACCACTCACTGAGCTTCTTAAAGTTTGAAGAGTTAGAGG 181
 Db 217 TTGTGGAAGAACCACTGACACCACTCACTGAGCTTCTTAAAGTTTGAAGAGTTAGAGC 276

Qy 182 ACTATACACTTTCTTTTGAAC 203
 Db 277 SCTATACACTTTCTTTTGAAC 298

RESULT 15

US-09-620-312D-61
 ; Sequence 61, Application US/09620312D
 ; Patent No. 6569662
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yungqing
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: John Tillinghast
 ; APPLICANT: Drmanac, Radolje T.
 ; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 ; FILE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 784CIP2B
 ; CURRENT APPLICATION NUMBER: US/09/620,312D
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 1105
 ; SOFTWARE: pt_fl_genes Version 1.0
 ; SEQ ID NO 61
 ; LENGTH: 2765
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (123)..(2291)
 US-09-620-312D-61

Query Match 4.6%; Score 68.8; DB 4; Length 2765;
 Best Local Similarity 62.4%; Pred. No. 1.7e-11;
 Matches 106; Conservative 1; Mismatches 63; Indels 0; Gaps 0;

Qy 697 GAATTTAGCTGTGAGGTATGCGGAGAGACATTTAGATCGCTTTTGTGATGATCCAC 756
 Db 939 GAGTTCGCTGCCAAGTGTGGGCCAGAGCTTTACACAGTCTTGTGTTCTCAAGGGCCAC 998

Qy 757 ATGAGACACACAAAGATTCTTTCACTTACGGGTGTACATGTGCGGAGAGAGATTCAAG 816
 Db 999 ATGCGTAGACCAAGGCCCTCTCTTCGATCATCGTGTCCGGGTGTGCGGCCGCTGCTTCAAG 1058

Qy 817 GAGCCTTGCTTTCTTTAAAAATCACATGCGGACRCATAATGGCAATCGG 866
 Db 1059 GAGCCTTGCTTTCTTTAAGAACCAATGAGGTGACGCCAGCAAGCTGG 1108

Search completed: June 29, 2005, 05:27:50
 Job time : 176.067 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2005, 01:06:08 ; Search time 589.497 Seconds
(without alignments)
15978.923 Million cell updates/sec

Title: US-08-731-499-3
Perfect score: 1507
Sequence: 1 GCAGTTCGTGGATTGACT.....GGCTTCGACTCCAGGTC 1507

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
21: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
22: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
23: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
24: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
25: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
26: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1504.8	99.9	1507	8	US-08-731-499-3
2	1387.2	92.1	10365	8	US-08-731-499-3
3	1387.2	92.1	36022	13	US-10-087-192-1708
4	1337.2	88.7	5632	15	US-10-177-293-505
5	1132.6	75.2	5252	13	US-10-087-192-1709
6	1106.2	73.4	3186	8	US-08-731-499-10
7	573.6	38.1	26345	13	US-10-087-192-1705

8	537	35.6	3016	13	US-10-087-192-1706	Sequence 1706, Ap
9	215.6	14.3	530	16	US-10-029-386-9711	Sequence 9711, Ap
10	156.4	10.4	469	13	US-10-040-739-520	Sequence 520, App
11	103.2	6.8	6033	17	US-10-172-118-1511	Sequence 1511, Ap
12	103.2	6.8	6033	18	US-10-342-887-1511	Sequence 1511, Ap
13	103.2	6.8	8156	14	US-10-074-475-93	Sequence 93, Appl
14	97	6.4	251	16	US-10-029-386-23411	Sequence 23411, A
15	70	4.6	70	20	US-10-758-307-110	Sequence 110, App
16	70	4.6	70	21	US-10-852-797-372	Sequence 372, App
17	68.8	4.6	1014	16	US-10-029-386-20910	Sequence 20910, A
18	68.8	4.6	1229	16	US-10-029-386-20193	Sequence 20193, A
19	68.8	4.6	2765	15	US-10-037-270-61	Sequence 61, Appl
20	68.8	4.6	2765	17	US-10-117-722-61	Sequence 61, Appl
21	68.8	4.6	3039	17	US-10-172-118-1644	Sequence 1644, Ap
22	68.8	4.6	3039	18	US-10-342-887-1644	Sequence 1644, Ap
23	68.8	4.6	3617	9	US-09-764-847-1394	Sequence 1394, Ap
24	68.8	4.6	3617	9	US-09-764-847-1394	Sequence 1394, Ap
25	68.8	4.6	3617	14	US-10-082-154-1394	Sequence 1394, Ap
26	68.8	4.6	3617	14	US-10-092-154-1396	Sequence 1396, Ap
27	59.2	3.9	508	16	US-10-029-386-7003	Sequence 7003, Ap
28	59.2	3.9	772	18	US-10-220-120-84	Sequence 84, Appl
29	59.2	3.9	2375	17	US-10-108-260A-896	Sequence 896, App
30	58.4	3.9	1595	9	US-09-764-864-99	Sequence 99, Appl
31	53.6	3.6	1584	9	US-09-764-864-550	Sequence 550, App
32	53.6	3.6	2014	9	US-09-764-864-100	Sequence 100, App
33	53.4	3.5	506	16	US-10-029-386-9132	Sequence 9132, Ap
34	53.4	3.5	1629	19	US-10-363-829-225	Sequence 225, App
35	53.4	3.5	2566	17	US-10-094-749-952	Sequence 952, App
36	53.4	3.5	3059	19	US-10-648-593-47	Sequence 47, Appl
37	53.4	3.5	3088	17	US-10-104-047-1016	Sequence 1016, Ap
38	52	3.5	594	16	US-10-029-386-25428	Sequence 25428, A
39	52	3.5	1912	19	US-10-479-435-45	Sequence 45, Appl
40	52	3.5	2072	17	US-10-094-749-728	Sequence 728, App
41	52	3.5	2261	17	US-10-104-047-1916	Sequence 1916, Ap
42	52	3.5	10519	9	US-09-764-847-1528	Sequence 1528, Ap
43	52	3.5	10519	14	US-10-082-154-1628	Sequence 1628, App
44	52	3.5	13337	9	US-09-764-846-312	Sequence 312, App
45	52	3.5	13337	9	US-09-764-847-1629	Sequence 1629, Ap

ALIGNMENTS

RESULT 1
US-08-731-499-3
; Sequence 3, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435


```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1507 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: 1..1507
; LOCATION: 1..1507
; OTHER INFORMATION: /note= "cDNA clone cc49 of 6-7kb
; OTHER INFORMATION: transcript with homology to C2H2 zinc
; OTHER INFORMATION: finger genes"
; US-08-731-499-3

Query Match 99.9%; Score 1504.8; DB 8; Length 1507;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGGTGCTGGGATGACTTCTGCTCAATGAAACACTCAATCAATGGAGACAAAGAG 60
DB 1 GCAGGTGCTGGGATGACTTCTGCTCAATGAAACACTCAATCAATGGAGACAAAGAG 60

QY 61 CACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCAATGGGAACCCCTGTATGCCCTTG 120
DB 61 CACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCAATGGGAACCCCTGTATGCCCTTG 120

QY 121 TTGTGGAAAGACAGTGACACCACTACAGTGCCTTCTAAAGTTCCGAAGTTAGAG 180
DB 121 TTGTGGAAAGACAGTGACACCACTACAGTGCCTTCTAAAGTTCCGAAGTTAGAG 180

QY 181 GACTATACACTTCTTTTGAACCTTTTATAATAATTTGCTCTGTTTGGAAACCCAGG 240
DB 181 GACTATACACTTCTTTTGAACCTTTTATAATAATTTGCTCTGTTTGGAAACCCAGG 240

QY 241 ACTGTTAGAGGTGAGTGACAGGCTTTACAGTGGCTTTAATCCAACTCCAGAAATGGCC 300
DB 241 ACTGTTAGAGGTGAGTGACAGGCTTTACAGTGGCTTTAATCCAACTCCAGAAATGGCC 300

QY 301 AACGGAACCTTTGAGATTATATGCAATCGAAAGTGCAGGAACATGCCAATCAATCCCT 360
DB 301 AACGGAACCTTTGAGATTATATGCAATCGAAAGTGCAGGAACATGCCAATCAATCCCT 360

QY 361 CTTAATGTACATGATGGCCAGAGTGATGGCAGCTCTCTTCCAGTCCGATGAGATG 420
DB 361 CTTAATGTACATGATGGCCAGAGTGATGGCAGCTCTCTTCCAGTCCGATGAGATG 420

QY 421 GAGATGCTTGTCAATGAAAGGCGCCNCTGTGTGTCAATTCGAGCTACACAAAGAAAAA 480
DB 421 GAGATGCTTGTCAATGAAAGGCGCCNCTGTGTGTCAATTCGAGCTACACAAAGAAAAA 480

QY 481 ATGTCAATCCGATCGAGGGGAATATGCCCTGGATTGCAATGCTTCTGACGACGACCTTC 540
DB 481 ATGTCAATCCGATCGAGGGGAATATGCCCTGGATTGCAATGCTTCTGACGACGACCTTC 540

QY 541 ACACATTCAGAAAGACTTTAATAAATGCTTTAATGCAACACCGGCTACCTCTGTGAA 600
DB 541 ACACATTCAGAAAGACTTTAATAAATGCTTTAATGCAACACCGGCTACCTCTGTGAA 600

QY 601 CCAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCCGCTTGTATATAAGTCAAGTGGGA 660
DB 601 CCAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCCGCTTGTATATAAGTCAAGTGGGA 660

```

```

QY 661 ACAGAACCTCCCAAGGAAAGAAATTTGCAAGGAAATGAATTTAGCTGTGAGGTATGTGG 720
DB 661 ACAGAACCTCCCAAGGAAAGAAATTTGCAAGGAAATGAATTTAGCTGTGAGGTATGTGG 720

QY 721 CAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTTTC 780
DB 721 CAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTTTC 780

QY 781 ACTTACGGGTGTAAATGTCGGGAAAGAGATTTCAAGGAGCCTTGTGTTCTTAAATATCAC 840
DB 781 ACTTACGGGTGTAAATGTCGGGAAAGAGATTTCAAGGAGCCTTGTGTTCTTAAATATCAC 840

QY 841 ATGGGACRCATATATGCAAAATCGGGGCCAGAACTGCAAGCAAGCTGCAAGGCTTGAGAGT 900
DB 841 ATGGGACRCATATATGCAAAATCGGGGCCAGAACTGCAAGCAAGCTGCAAGGCTTGAGAGT 900

QY 901 AGTCCAGCAACGATCAACGAGGTCTGTCAGGTGCAACGCGGCGGAGAGCATCTCTCTCT 960
DB 901 AGTCCAGCAACGATCAACGAGGTCTGTCAGGTGCAACGCGGCGGAGAGCATCTCTCTCT 960

QY 961 TGCAAAATCTGATGTTGTTGCTTCTATTTCCAAATAAAGAAAGTCTAATTTAGACAC 1020
DB 961 TGCAAAATCTGATGTTGTTGCTTCTATTTCCAAATAAAGAAAGTCTAATTTAGACAC 1020

QY 1021 CGCAAGTGCACACCAAAATGCTTTTCTGTCAGGAGGAGCTTCTGTCAGGTGTTTCAACTGAGACCAAAA 1140
DB 1021 CGCAAGTGCACACCAAAATGCTTTTCTGTCAGGAGGAGCTTCTGTCAGGTGTTTCAACTGAGACCAAAA 1140

QY 1141 TCTCACTCTGAAAACGGGGAAGAGCTGTGATGTCATCCCTCAGCTCGATCCGTTTCAAC 1200
DB 1141 TCTCACTCTGAAAACGGGGAAGAGCTGTGATGTCATCCCTCAGCTCGATCCGTTTCAAC 1200

QY 1201 ACCTTCCAGGCTTGCACTGCTACCAAGAAAGTTGCCATTTGCCAAGAGTGAAG 1260
DB 1201 ACCTTCCAGGCTTGCACTGCTACCAAGAAAGTTGCCATTTGCCAAGAGTGAAG 1260

QY 1261 GAAATTTGGGGAAGAGGAGGAGCAACCAACGATTCGAGTTCGAGAGAGGAGCTTGA 1320
DB 1261 GAAATTTGGGGAAGAGGAGGAGCAACCAACGATTCGAGTTCGAGAGAGGAGCTTGA 1320

QY 1321 GAAACAAATAGAACCAATTTGTCAGGCTCTCGGAGAGAAAGAGTGGCAACACTCC 1380
DB 1321 GAAACAAATAGAACCAATTTGTCAGGCTCTCGGAGAGAAAGAGTGGCAACACTCC 1380

QY 1381 CACGGCAAGCGCCCTCCGTCGACGCGATCCCAAGTTACCCAGTAGCAAGAGAGGCC 1440
DB 1381 CACGGCAAGCGCCCTCCGTCGACGCGATCCCAAGTTACCCAGTAGCAAGAGAGGCC 1440

QY 1441 ACTCATCTGCTCGAGTGGCAAGCTTTTCAGAACTTACCACAGCTGGTCTTGCACCTCC 1500
DB 1441 ACTCATCTGCTCGAGTGGCAAGCTTTTCAGAACTTACCACAGCTGGTCTTGCACCTCC 1500

QY 1501 AGGTCC 1507
DB 1501 AGGTCC 1507

```

```

RESULT 2
US-08-731-499-9
; Sequence 9, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR

```

;; TITLE OF INVENTION: USES
;; NUMBER OF SEQUENCES: 44
;; CORRESPONDENCE ADDRESS: 44
;; ADDRESSEE: Townsend and Townsend and Crew
;; STREET: Two Embarcadero Center, 8th Floor
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94111-3834
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/731,499
;; FILING DATE: 16-OCT-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/680,395
;; FILING DATE: 15-JUL-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hunter, Tom
;; REGISTRATION NUMBER: 38,498
;; REFERENCE/DOCKET NUMBER: 23070-068910
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10365 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: -
;; LOCATION: 1..10365
;; OTHER INFORMATION: /note= "Genomic Sequence Encoding
;; OTHER INFORMATION: ZABC1"
US-08-731-499-9

Query Match 92.1%; Score 1387.2; DB 8; Length 10365;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY	2	CAGGTTCCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAGAGC	61
DB	2424	CAGGTTCCTGGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAGAGG	2483
QY	62	ACTAATGCTTTTGCTGATTTCATATTGAAATCGAGGCATTGGGAAACCTGTATGCTTGT	121
DB	2484	ACTAATGCTTTTGCTGATTTCATATTGAAATCGAGGCATTGGGAAACCTGTATGCTTGT	2543
QY	122	TTGTGAAAGAACCGATGACACATCAGTCTTCTTAAAGTTGCAAGAAAGTTAGAGG	181
DB	2544	TTGTGAAAGAACCGATGACACATCAGTCTTCTTAAAGTTGCAAGAAAGTTAGAGG	2603
QY	182	ACTATACACTTCTTTTGAACCTTTTATATAATAATTGCTCTGG-TTTTGAACCCAGG	240
DB	2604	ACTATACACTTCTTTTGAACCTTTTATATAATAATTGCTCTGGTTTGAACCCAGG	2663
QY	241	ACTGTTTGAAG-GGGTGAAGTGAAGTCTTAC-AGTGGCCCTTAATCCAACTCCAGAAATGTC	298
DB	2664	GCTGTTAGAGGGTGAAGTGAAGTCTTACAGTGGCCCTTATCCAACTCCAGAAATGTC	2723
QY	299	CCAAAGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC	358
DB	2724	CCAAAGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC	2783
QY	359	CTCTTAATGTACATGAT-GGGCAAGAGTGATGGCAGCTCTCTTGGCCAGTCCGATGGAG	417
DB	2784	CTCTTAATGTACATGATGGGCAAGAGTGATGGCAGCTCTCTTGGCCAGTCCGATGGAG	2843

QY	418	ATGGA-GATGCTTGTCAATGAAAGGGCCNCTGTTGTCAATTCGAGCTACACAAAGAA	476
DB	2844	ATGGAAGGATGCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCGAGCTACACAAAGAA	2903
QY	477	AAAAATGTCATCCGAATCGAGGGGAATATGCCCTTGATTCGATTCGAGCCAGAC	536
DB	2904	AAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGATTCGATTCGAGCCAGAC	2962
QY	537	CTTCACACATTCAGAAAGACCTTAATAAATCATGCTTAATGCAACACCGCCCTACCCCTCTG	596
DB	2963	CTTCACACATTCAGAAAGACCTTAATAAATCATGCTTAATGCAACACCGCCCTACCCCTCTG	3022
QY	597	TGAACACAGCTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT	656
DB	3023	TGAACACAGCTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT	3082
QY	657	GCGAAGACACCTCCNAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATG	716
DB	3083	GCGAAGACACCTCCNAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATG	3142
QY	717	TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAAACACACAAAGATTTC	776
DB	3143	TGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCACATGAGAAACACACAAAGATTTC	3202
QY	777	TTTCACTTACGGGTGTAAATGTGCGGAAGAAAGATTCAAGGAGCCTTGGTTCTTTAAAAA	836
DB	3203	TTTCACTTACGGGTGTAAATGTGCGGAAGAAAGATTCAAGGAGCCTTGGTTCTTTAAAAA	3262
QY	837	TCATATGCGGACRCAATATGCGCAATCGGGGCCAGAGCAAACTGCAGCAGGCTTGA	896
DB	3263	TCATATGCGGACRCAATATGCGCAATCGGGGCCAGAGCAAACTGCAGCAGGCTTGA	3322
QY	897	GAGTAGTCAGCAACGATCAACGAGTCTGTCAGAGTGACGCGCGCGAGAGCATCTCCTC	956
DB	3323	GAGTAGTCAGCAACGATCAACGAGTCTGTCAGAGTGACGCGCGCGAGAGCATCTCCTC	3382
QY	957	TCCTTGCAAAATCTGCATGTTTGGTTCCTATTTCAAATAAAGAAAGTCTAATGA	1016
DB	3383	TCCTTGCAAAATCTGCATGTTTGGTTCCTATTTCAAATAAAGAAAGTCTAATGA	3442
QY	1017	GCACGCAAGGTGACACCAAAAAAATCTGCTTCGTACAGCAGCGCGCGAGAGCTC	1076
DB	3443	GCACGCAAGGTGACACCAAAAAAATCTGCTTCGTACAGCAGCGCGCGAGAGCTC	3502
QY	1077	TCACAAGGAGAAATGCGTCTCGAGGAGAGCTTCTGCAAGTTGTTCAACTTGAGACC	1136
DB	3503	TCACAAGGAGAAATGCGTCTCGAGGAGAGCTTCTGCAAGTTGTTCAACTTGAGACC	3562
QY	1137	AAAACTCACCCTGAAAACGGGGAAGAGCCTGTGAGATGCAATCCCTCAGCTCGATCCGTT	1196
DB	3563	AAAACTCACCCTGAAAACGGGGAAGAGCCTGTGAGATGCAATCCCTCAGCTCGATCCGTT	3622
QY	1197	CACCACTTCCAGGCTTGGCAKCTGCTTACCAAGAGAAAGTTGCCATTTGCAAGAGT	1256
DB	3623	CACCACTTCCAGGCTTGGCAKCTGCTTACCAAGAGAAAGTTGCCATTTGCAAGAGT	3682
QY	1257	GAAGCAATTTGGGCAAGAGGAGCAGCAGCAACGATTCGAGTTCGAGAGAGGAGCT	1316
DB	3683	GAAGCAATTTGGGCAAGAGGAGCAGCAGCAGCAACGATTCGAGTTCGAGAGAGGAGCT	3742
QY	1317	TGGAGAAACAAATAGAACCAATTTGTGAGGCTCTCGCAAGAGAAAGAGTGCAGAAACA	1376
DB	3743	TGGAGAAACAAATAGAACCAATTTGTGAGGCTCTCGCAAGAGAAAGAGTGCAGAAACA	3802
QY	1377	CTCCACGCGGAGCGCCCTCGTGGAACGCGGATCCCAAGTTACCAGTAGCAAGAGAA	1436
DB	3803	CTCCACGCGGAGCGCCCTCGTGGAACGCGGATCCCAAGTTACCAGTAGCAAGAGAA	3862
QY	1437	GCCCACTCACTGCTCCGAGTGGGCAAGAGCTTTAGAACCTTACCAAGCTGCTTGA	1496
DB	3863	GCCCACTCACTGCTCCGAGTGGGCAAGAGCTTTAGAACCTTACCAAGCTGCTTGA	3922

QY 1497 CTCACGGGTCC 1507
 DB 3923 CTCACGGGTCC 3933

RESULT 3
 US-10-087-192-1708
 ; Sequence 1708, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; FILE OF INVENTION: CANCER
 ; FILE REFERENCE: 52945200122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; PRIOR FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1708
 ; LENGTH: 36022
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-087-192-1708

Query Match 92.1%; Score 1387.2; DB 13; Length 36022;
 Best Local Similarity 97.3%; Pred. No. 0;
 Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 2 CAGGTTCGTGGATTGACTTCTTGCTCAATTGAAACACTCAATTCATGAGACAAAGAGC 61
 DB 9951 CAGGTTCGTGGATTGACTTCTTGCTCAATTGAAACACTCAATTCATGAGACAAAGAG 10010

QY 62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTTGGACCCCTGTATGCTTGT 121
 DB 10011 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTTGGACCCCTGTATGCTTGT 10070

QY 122 TTGTGAAAGAACCAAGTGACACCATCACTGAGCTTCTTAAAGTTTGAAGAGTTAGAGG 181
 DB 10071 TTGTGAAAGAACCAAGTGACACCATCACTGAGCTTCTTAAAGTTTGAAGAGTTAGAGG 10130

QY 182 ACTATACACTTTCTTTTGAACTTTTATATAATTAATTTGCTCTGG-TTTTGGAACCCAGG 240
 DB 10131 ACTATACACTTTCTTTTGAACTTTTATATAATTAATTTGCTCTGGTTTTTGAACCCAGG 10190

QY 241 ACTGTTTGA-GGGTGAGTGACAGGTCTTAC-AGTGGCCTTAAATCCAACTCCAGAAATTC 298
 DB 10191 GCTGTTAGAGGGTGAGTGACAAAGTCTTACAAAGTGGCCCTTATTCCACTCCAGAAATTC 10250

QY 299 CCAACGGAACCTTGAGATTATATGCAATCGAAAGTGACAGAAACATGCCAACTCAATCC 358
 DB 10251 CCAACGGAACCTTGAGATTATATGCAATCGAAAGTGACAGAAACATGCCAACTCAATCC 10310

QY 359 CTCCTTAATGTACATGGAT-GGCAAGAGTGATGAGCAGCTCTTTGCGCAGTCCGATGAG 417
 DB 10311 CTCCTTAATGTACATGGATGAGCAGCTCTTTGCGCAGTCCGATGAG 10370

QY 418 ATGGA-GATGCTTGTCAATGAAGGGCCCTGTTGTCAATTCGAGCTACACAAAGA 476
 DB 10371 ATGGAAGATGCTTGTCAATGAAGGGCCCTGTTGTCAATTCGAGCTACACAAAGA 10430

QY 477 AAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCAATGTTCTGAGCCAGAC 536
 DB 10431 AAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCAATGTTCTGAGCCAGAC 10489

QY 537 CTTACACATTCAGAAACCTTAAATAACATGCTTTAATGCAACACCGGCCCTACCTCTG 596
 DB 10490 CTTACACATTCAGAAACCTTAAATAACATGCTTTAATGCAACACCGGCCCTACCTCTG 10549

QY 597 TGAACACGACGTTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGT 656
 DB 10550 TGAACACGACGTTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGT 10609

QY 657 GCGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATG 716
 DB 10610 GCGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATG 10669

QY 717 TGGGACAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTC 776
 DB 10670 TGGGACAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTC 10729

QY 777 TTTCACTTACGGGTGTAACATGTGCGGAAGAAAGATTCAAGGAGCCTTGTTCTTTAAAAA 836
 DB 10730 TTTCACTTACGGGTGTAACATGTGCGGAAGAAAGATTCAAGGAGCCTTGTTCTTTAAAAA 10789

QY 837 TCACATGCGGACRCATATATGGAATCGGGGCCAGAAAGCAAACTGACGAGGCTTGA 896
 DB 10790 TCACATGCGGACACATATATGGAATCGGGGCCAGAAAGCAAACTGACGAGGCTTGA 10849

QY 897 GAGTAGTCAGAACGATCAACGAGGTGCTCAGGTGACGGGGCCGAGAGCATCTCTC 956
 DB 10850 GAGTAGTCAGAACGATCAACGAGGTGCTCAGGTGACGGGGCCGAGAGCATCTCTC 10909

QY 957 TCCTTGCAAAATCTGCATGTTTGGCTTCTTATTTCCAAATTAAGAAAGTCTAATGA 1016
 DB 10910 TCCTTGCAAAATCTGCATGTTTGGCTTCTTATTTCCAAATTAAGAAAGTCTAATGA 10969

QY 1017 GCACCGCAAGGTGCACACCAAAAAAATCTGCTTTCGTTACCAAGCAGCGCGCAGACAGCTC 1076
 DB 10970 GCACCGCAAGGTGCACACCAAAAAAATCTGCTTTCGTTACCAAGCAGCGCGCAGACAGCTC 11029

QY 1077 TCCACAGGAGGATGCGCTCTCGAGGAGGACCTTCTGACAGTTGTTCAACTTGAGACC 1136
 DB 11030 TCCACAGGAGGATGCGCTCTCGAGGAGGACCTTCTGACAGTTGTTCAACTTGAGACC 11089

QY 1137 AAAATCTCACCTCGAAACGGGGAAGAGCCCTGTGATGATGATCCCTCAGCTCGATCCGTT 1196
 DB 11090 AAAATCTCACCTGAAACGGGGAAGAGCCCTGTGATGATGATCCCTCAGCTCGATCCGTT 11149

QY 1197 CACCACTTTCAGGCTTGGCAKCTGGCTTACAAAGGAAGTTCATTTGCCAAGAGT 1256
 DB 11150 CACCACTTTCAGGCTTGGCAKCTGGCTTACAAAGGAAGTTCATTTGCCAAGAGT 11209

QY 1257 GAAAGATTTGGGCAAGAGGAGCAGCAGACGACGATTCGAGTTCGAGAGGAGCT 1316
 DB 11210 GAAAGATTTGGGCAAGAGGAGCAGCAGACGACGATTCGAGTTCGAGAGGAGCT 11269

QY 1317 TGGAGAAACAAATAAGAACCACTTGTGAGGCTCTTCGCAAGAGAAAGAGAGTCAAAACA 1376
 DB 11270 TGGAGAAACAAATAAGAGGAGCTTGTGAGGCTCTTCGCAAGAGAAAGAGAGTCAAAACA 11329

QY 1377 CTCCTCACGGAAGCGCCCTCCGTGGAGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 1436
 DB 11330 CTCCTCACGGAAGCGCCCTCCGTGGAGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 11389

QY 1437 GCCCACTCAGCTCTCGAGTGGGCAAGAGCTTTCAGAACCTTACCAAGCTGCTTGA 1496
 DB 11390 GCCCACTCAGCTCTCGAGTGGGCAAGAGCTTTCAGAACCTTACCAAGCTGCTTGA 11449

QY 1497 CTCACGGGTCC 1507
 DB 11450 CTCACGGGTCC 11460

RESULT 4
 US-10-177-293-505
 ; Sequence 505, Application US/10177293
 ; Publication No. US20030124128A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lillie, James
 ; APPLICANT: Glatt, Karen
 ; APPLICANT: Zhao, Xumei


```

US-10-087-192-1709
; Sequence 1709, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1709
; LENGTH: 5252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1709

Query Match      75.2%; Score 1132.6; DB 13; Length 5252;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1180; Conservative 3; Mismatches 28; Indels 3; Gaps 3;

QY 296 TGCCCAACGGAACTTGTAGATTATATGCAATGGAAGTGCAGAGAAATGCGCAACTCAA 355
Db 1 TGCCCAACGGAACTTGTAGATTATATGCAATGGAAGTGCAGAGAAATGCGCAACTCAA 60

QY 356 TCCCTCTTAATGTACATCGAT-GGCCAAGAGTGTGGCAGCTCTCTTGCAGTCCGATG 414
Db 61 TCCCTCTTAATGTACATCGATGGCGCAGAGTGTGGCAGCTCTCTTGGCAGTCCGATG 120

QY 415 GAGATGGA-GATGCCCTTGTCAATGAAGGGCCNCTGTGTCAATTCGAGCTACACAA 473
Db 121 GAGATGGAAGTGCCTTGTCAATGAAGGGCCGCTGTGTTCATTCGAGCTACACAA 180

QY 474 GAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGAGCA 533
Db 181 GAAAAAATGTG-C-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGAGCA 239

QY 534 GACCTTCACATTCAGAGACCTTAATAACATGCTTAATGCAACACCGCCCTACCCCT 593
Db 240 GACCTTCACATTCAGAGACCTTAATAACATGCTTAATGCAACACCGCCCTACCCCT 299

QY 594 CTGTGAACCCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCA 653
Db 300 CTGTGAACCCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCA 359

QY 654 AGTGGGAACAGAACCTCCAGAGAAAGAAATGCAAGAAATGAATTTAGTGTGAGGT 713
Db 360 AGTGGGAACAGAACCTCCAGAGAAAGAAATGCAAGAAATGAATTTAGTGTGAGGT 419

QY 714 ATGTGGGAGACATTTAGAGTGCCTTTTGTGATGTTCCAGATCCACATGAGAACACAAAGA 773
Db 420 ATGTGGGAGACATTTAGAGTGCCTTTTGTGATGTTCCAGATCCACATGAGAACACAAAGA 479

QY 774 TTCTTTTCACTTACGGGTGTAAATGTCGGGAAGAAAGATTCAAGGAGCCCTTGGTTTCTTAA 833
Db 480 TTCTTTTCACTTACGGGTGTAAATGTCGGGAAGAAAGATTCAAGGAGCCCTTGGTTTCTTAA 539

QY 834 AAATCAGATGGGACRCATATATGGAATCGGGGCCAGAGCAAACTGCAGCAGGCTT 893
Db 540 AAATCAGATGGGACRCATATATGGAATCGGGGCCAGAGCAAACTGCAGCAGGCTT 599

QY 894 GGAGAGTAGTCAGCAACGATCAACAGGTCTCTCCAGGTGCACGGCCGCGAGAGCATCTC 953
Db 600 GGAGAGTAGTCAGCAACGATCAACAGGTCTCTCCAGGTGCACGGCCGCGAGAGCATCTC 659

QY 954 CTCTCTTTGCAAAATCTGCATGGTTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAAT 1013
Db 1013 CTCTCTTTGCAAAATCTGCATGGTTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAAT 719

```

```

RESULT 6
US-08-731-499-10
; Sequence 10, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:

```

```
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3186 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..3186
; OTHER INFORMATION: /note= "ZABC1 Open Reading Frame"
US-08-731-499-10

Query Match      73.4%; Score 1106.2; DB 8; Length 3186;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1150; Conservative 9; Mismatches 28; Indels 3; Gaps 3;

QY    320 ATGCAATCGAAAGTCACAGGAACAATGCCAACTCCAATCCCTCTTAATGTACATGGAT-GG 378
Db    |||
QY    1   ATGCAATCGAAAGTCACAGGAACAATGCCAACTCCAATCCCTCTTAATGTACATGGATGGG 60
Db    |||

QY    379 CCAAGAGTGATTGGCAGCTCTTTGCCAGTCCGATGGAGATGGA-GATGCCTTGTCTCAATG 437
Db    |||
QY    61  CCAGAGTGATTGGCAGCTCTTTGGCAGTCCGATGGAGATGGAGGATGCCTTGTCTAANTG 120
Db    |||

QY    438 AAAGGGCCNCNTGTTGTCAATTC CGAGCTACACAAAGAAAAAATGTCAATCCGAATCGA 497
Db    |||
QY    121 AAAGGGACCCTGTTGTTCCATTCCGAGCTACACAAGAAAAAATGTC-ATCCAATTCGA 179
Db    |||

QY    498 GGGGAATATGCCCTTGGATTGCATGTTCTGCAGCAGACCTTCACATTTCAGAAGACCT 557
Db    |||
QY    180 GGGGTATATGCCCTTGGATTGCATGTTCTGCAGCAGACCTTCACATTTCAGAAGACCT 239
Db    |||

QY    558 TAATAAACATGCTTTAATGCAACACCGGCCCTACCCTCTGTGAAACGACGAGTCTTTCGGGT 617
Db    |||
QY    240 TAATAAACATGCTTTAATGCAACACCGGCCCTACCCTCTGTGAAACGACGAGTCTTTCGGGT 299
Db    |||

QY    618 TGAAGCAGAGTATCTCAGTCGCGTTGATAAAAAGTCAAAGTCGGAACGAACTCCCAAGGA 677
Db    |||
QY    300 TGAAGCAGAGTATCTCAGTCGCGTTGATAAAAAGTCAAAGTCGGAACGAACTCCCAAGGA 359
Db    |||

QY    678 AAGAATTGCAAGAAAAATGAATTTAGCTGTGAGGTTATGTGGGCAGACATTTTAGAGTCGC 737
Db    |||
QY    360 AAGAATTGCAAGAAAAATGAATTTAGCTGTGAGGTTATGTGGGCAGACATTTTAGAGTCGC 419
Db    |||

QY    738 TTTTGTGTTTGAGATCCACATGGAACACACAAGAATTCCTTCACTTACCGGTGTATACAT 797
Db    |||
QY    420 TTTTGTGTTTGAGATCCACATGGAACACACAAGAATTCCTTCACTTACCGGTGTATACAT 479
Db    |||

QY    798 GTCCGAAGAAGATTCAAGGAGCCTTGCTTTCTTAAAAATACATGCGGACRCAATAATGG 857
Db    |||
QY    480 GTCCGNAGAGMTTSRSSAGCCTTGTTTTCTTAAAAATCAATGCGGACACATANTGG 539
Db    |||

QY    858 CAATCGGGGGCCAGAAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAA 917
Db    |||
QY    540 CAATCGGGGGCCAGAAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAA 599
Db    |||

QY    918 CGAGGTCGTCCAGGTGCAACGGGCCGAGAGCATCTCTCTCTTGGCAAAATCTGCATGGT 977
Db    |||
QY    600 CGAGGTCGTCCAGGTGCAACGGGCCGAGAGCATCTCTCTCTCTTACAAAAATCTGCATGGT 659
Db    |||

QY    978 TTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTTAATTGAGCACCGCAAGGTGCACACAA 1037
Db    |||
QY    660 TTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTTAATTGAGCACCGCAAGGTGCACACAA 719
Db    |||

QY    1038 AAAAATGCTTTTCGGTACCGAGCGCGCAGACAGACTCTTCCAAGGAGGAATTCGCCTC 1097
```


Db 9839 AATATACTCGTCCCTCTGACCTT-----AGAGGTGCTCTGATTTTGACACCCAGG 9889
Qy 242 CTGTTAGAGGTGAG---TGACAGTCTTACAGTGGCCTTAATCAACTCCAGAAATTGC 298
Db 9890 CTCCTTAGAGGTCTGAGTGGTGATATCTCACAGGCGGCTTATTCTAACTGCAGAGATCC 9949
Qy 299 CCAACGGAACTTTGAGATTATATGCAATCGA-----AAGTGACAGGAAACATGCCAACT 352
Db 9950 TCAGCAGATAGTGGATTACCCATCTCTGACACCTGAAAGTGGTGAGCAGATGCCGACC 10009
Qy 353 CAATCCCTCTTAATGATACATGA-TGGCCAAAGATGATTGGCAGCTCTCTTTGCCAGTCCG 411
Db 10010 CAGTCCCTCTCGTGATACATGACGCGCGGGAAGTCTCAGCAGCTCTCTAGGCTCCAG 10069
Qy 412 ATGGAGATGGAGATGCCCTTGTCAATGAAGGGCCNCTGTGTGTCATTTCCGAGCTACACA 471
Db 10070 ATGGAGGTGGATGATGCTGTGCCCCATGAAGGGCCGTTGGCAGTCCCTTCGAGTCTCT 10129
Qy 472 AAGAAAAAATGTCATCCGAATCGAGGGGAATATGCCCTTGGATTGCAATGTTCTGACGC 531
Db 10130 CAGGAGAGAGCATGGCCGTGGCAGAGGGCCACATGCCCTGGATTGCAATGTTCTGACGC 10189
Qy 532 CAGACCTTCACACATTCAGAAAGACCTTAATAACATGTTCTTAATCAACACCGGCTTACC 591
Db 10190 CAGGTTCTCTCAGCGCGAGGATCTCAGTCAGCAGTGTCTGTCAGCACCGGCCACCC 10249
Qy 592 CTCTGTGAACCGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGT 651
Db 10250 CTCTCGAGCCAGCTGTCTCTGCTGTGGAGCGGAGTACCTAAAGTCCCTTTGATAAGCT 10309
Qy 652 CAAGTGGCAACAGAACCTCCCAAGAAAGAAATTCGAAGGA---AAATGAATTTAGTGT 708
Db 10310 CTGGAGCCACAGAGCCAGCATTTGGAAGAGAGTGCGGAAGACCCGAGAGGTTGAGCTGT 10369
Qy 709 GAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATGAGAACAC 768
Db 10370 GATGTGTGGGCAGACATTCACAGTGTCTTGTATGTTGAGAGCCACATGAAGAAGCAT 10429
Qy 769 AAGATTTCTTCACATTACGGGTGTAACATGTGCGGAAGAGATTTCAAGGAGCCTTGGTT 828
Db 10430 AAGGACTCTTCACGTATGCGTGTGAGCATGTGCGGGAGGAGATTCAGGAGCCGTGGTTC 10489
Qy 829 CTTAAAAATCACATCGGACRCATATGCAAAATCGGGGGCCAGAACGAACTGCAGCAA 888
Db 10490 CTGAAGAACCATGCGGACACACATGCAAGTCTGSCACAGGAGCAAGCTTCAGCAA 10549
Qy 889 GGTGGAGAGTAGTCCAGCAACGATCAACGAGGTGTCAGGTGCAAGCGGCCGAGAGC 948
Db 10550 GGCATGGAG---AGTCCAGTCACCATCAATGAAGTGTCCAGCCGACGCCCCCTGGAGC 10606
Qy 949 ATCTCTCTCTCTTGCAAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAGAAAGT 1008
Db 10607 ATCTCACGCCCTACAAGATCTGCATGTTCTGCGGCTTCTCTTCCCAATAAGCAGAGC 10666
Qy 1009 CTAATTGAGCCGCAAGGTGCACACCAAAACTGCTTTTCGGTACCAGCAGCGCGCAG 1068
Db 10667 CTATTGAGCAGCAGAGGTTCAAGCCAAAGAACTGTCCCGAGTGCAGCAACGTTGCC 10726
Qy 1069 ACAGATCTCCCAAGGAGGAATGCCGTCTCGAGGAGGAGTCTCCTCAGTGTGTTCAAC 1128
Db 10727 CTTGATGACCCAGAGAGGAACCCACGTCCTCCGAGGGAAGAGTTGCTGCAGTTTTTGAAC 10786
Qy 1129 TTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCCTGTGAGATGATCCCTCAGCTC 1188
Db 10787 TTGAGACCCAGATCAACTGCAAGGTAGTACAGTGAAGCCCATGACCTGCATACCTCAGCTT 10846
Qy 1189 GATCCGTTTCAACACCTTCCAGCTTTGGCAKCTGGCTTACCAAGGAAGATGTTGCCATTTGC 1248
Db 10847 GACCGGTTTCAACACCTTACAGGATGAGCAGTTGGCTTACCAAGGAAGAGTGGCGGTTGCC 10906
Qy 1249 C---AAGAGTGAAGGAATTTGGGCAAGAGGAGCAACGACAAAGAGTTCGAGTTCC 1305
Db 10907 CAGGAAGAGGTGAAGAGTTCAGGCCAAGAGGAAGCAGACAATGACGACTCATGCTCA 10966

Qy 1306 GAGAGGAGCTTGGAG----- 1321
Db 10967 GAGAAAGAGGAAGAACTTAGGGGAATAATATGGGTTGGGGTAAGCGGAAGGGTCTGGAAGAATGCC 11026
Qy 1322 -AAACAAATAAGAAACCAATTGTGCGAGGCCTCTCGCAAGAGAAAGAGAGTGAAGCACTCC 1380
Db 11027 AAACNAGTAAAGCAGTTGTCCAGGTCTCTCCDAGACAAGAGAGCCTAGACATGCT 11086
Qy 1381 CACGGCGAAGCGCCCTCCGTGGAGCGCGATCCCAAGTTACCCAGTACGAGAGAGAGCC 1440
Db 11087 AATAGTGAAGTGCCTTCTGGGGATAGTACCCCAAGTTGTCCAGTAGCAAGAGAGGCC 11146
Qy 1441 ACTCACTCTCCGAGTGGCGCAAGCTTTTCAAGACCTTACCACAGCTGCTTTCACATCC 1500
Db 11147 ACGCACTCTTCTGAGTGCAGCAAGCTTTTCAAGACATACCACAGCTCGTCTGCACTCG 11206
Qy 1501 AGGGTCC 1507
Db 11207 AGGGTGC 11213

RESULT 8
US-10-087-192-1706
; Sequence 1706, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1706
; LENGTH: 3016
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-1706

Query Match 35.6%; Score 537; DB 13; Length 3016;
Best Local Similarity 68.6%; Pred. No. 1.8e-157;
Matches 834; Conservative 3; Mismatches 324; Indels 55; Gaps 5;

Qy 344 ATGCCAACTCAATCCCTCTTAATGTACATGGA-TGGCCAAAGAGTGAATGGCAGCTCTCTT 402
Db 1 ATGCGGACCCAGTCCCTCTCGTGTACATGGAACGCGCCGGAAGTCTTCAGCAGCTCTCTTA 60
Qy 403 GCCAGTCCGATGGAGATGGAGATGCTTGTCAATGAAGGGCCNCTGTGTCAATTCG 462
Db 61 GGCTCCAGATGGAGGTGGATGCTGTGCCATAAAGGGCCGCGTGGCAGTCCCCCTC 120
Qy 463 AGCTACACAAAGAAATAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCAATG 522
Db 121 CGAGCTGTCTAGGAGAGAGCATGCGCCGTGGCAGAGGGGCCACATGCCCTTGGATTGCAATG 180
Qy 523 TTCTGCGAGCAGACTTTCACACATTCAGAACCTTAATAACATGTCTTAATGCAACAC 582
Db 181 TTCTGCGAGCCAGGTCTTCTCTCAGCGGAGGATCTCAGTCAGCAGCTGCTGTCGAGCAC 240
Qy 583 CGGCTACCTCTGTGAAACCCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTT 642
Db 241 CGGCCCACTCTCGAGCCAGCTGTCTCTGCTGTGGAGCGCGAGTACTTAAGTCCCTT 300
Qy 643 GATAAAGTCAAGTGGCAACAGAACCTCCCAAGGAAGAAATTCGAAGA---AAATGA 699
Db 301 GATAAAGTCTGGAGCCCAACAGAGCCAGCATTTGGAGAGAGTGGCGGAAGACCCCGAGGAG 360


```
QY 700 TTATGCTGTGAGTATGTGGGAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATG 759
D 361 TTGAGCTGTGATGTGTGGGAGACATTTCCAGTGGCTTTTGTATGTTGAGAGCCACATG 420
QY 760 AGAACAACAAGATCTTTCTTACTTACGGGTGTAAATGTCGGAGAGATTCAGAG 819
D 421 AAGAAGCATAGGACTCTTCTCATGATGGTGCAGCATGTGGGGAGGAGATTCAGAG 480
QY 820 CTTTGTTTCTTAAATACATGCGACRCAATATGCAAACTCGGGGCCAGAGCAA 879
D 481 CGTGTCTTGNAGNACCATGCGACACACATGCGAGTCTGGCACAGGAGCAAG 540
QY 880 CTGAGCAAGGTTGAGATGTAGTCAGCAACGATCAACAGAGTGTGTGAGTGCACGCG 939
D 541 CTTTCAAGCAAGGATGAG---AGTCAGTCAACATCAATGAAGTGTGTGAGCGGACGCG 597
QY 940 GCGGAGAGCATCTCTCTCTTGCATAAATCTGCATGTTTGTGGTCTCTATTTCCAAT 999
D 598 CTGGGAGCATCTCACGCGCTTACAGATCTGCATGTTGTGGGTCTCTTCTTCCCAAT 657
QY 1000 AAGAAGTCTAATTCAGCAGCGCAAGGTGCACACCAAACTGCTTTCGGTACCAGC 1059
D 658 AAGCAGAGCTCATTTGAGCAGAGGATTTACGCCAAGAACTGTCCCGAGTGCAGC 717
QY 1060 AGCGGCGAGACAGATCTTCCACAGGAGGANTGCCGTCTTCAGGGAGGACTTCTGTGCA 1119
D 718 AACGTTGCCCTGTATGATCACCAGGAGAAACCAAGTCCCGGAGGAGTGTGTGCA 777
QY 1120 TTGTTCACTTGAACCAAAATCTACCTGTAACCGGGAGAGCTGTGATGTCATC 1179
D 778 TTTTGAATTTGAGACCCCAAGATCAATGTCAGGTAGTACAGTGAAGCCCATGACCTG 837
QY 1180 CTTAGCTGATGTCGTTTACACCTTCCAGGCTTGCACCTGCTACCAAGGAAAGT 1239
D 838 CTTAGCTTGAACCCGTTTACACCTTACCAGGATGCGAGTGTGGTACCAAGGAAAGT 897
QY 1240 GCCATTTGCC---AAGAAGTGAAGAAATTTGGGGCAAGAGGAGGAGCACCACACGACG 1296
D 898 GCGGTTGCCAGGAGAGGTTGAAGAGTCAAGGACAGGAGGAGGAGCAACATGACGAC 957
QY 1297 TCGAGTTCCGAGAGGAGTTCGAG----- 1321
D 958 TCATGCTCAGAGAAAGAGGAACTAGGGGAAATATGGGTTGGGGTAAGGCGGAGGTC 1017
QY 1322 -----AAACAAATGAAGAACCATTTGTGAGGCTCTCCAGAGCAAGAGAGTGC 1371
D 1018 GGAAGTCCAAACAAAGTAAAGCAAGTGTGTCCAGGTCTCTCCAGAGCAAGAGAGCCT 1077
QY 1372 AAACACTCCAGCGGAGGCGCTTCCGTGAGAGGAGTCCCAAGTTACCCAGTAGCAAG 1431
D 1078 AGACATGCTAATAGTGAAGTGTCTTGGGATAGTGACCCCAAGTTGTCCAGTAGCAAG 1137
QY 1432 GAGAAGCCCACTGCTGTCCAGTGGCGGCAAGCTTTAGAACCTTACACAGCTGTGTC 1491
D 1138 GAGAAGCCCACTGCTGTGAGTGCAGAAAGCTTCCAGGACATACCAAGCTGTGTC 1197
QY 1492 TTGCACTCCAGGCTCC 1507
D 1198 CTGCACTCGAGGTC 1213
```

```
RESULT 9
US-10-029-386-9711/c
; Sequence 9711, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
```

```
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9711
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: SWISSPROT HIT: O75362, EVALUE 2.00e-67
; OTHER INFORMATION: NT HIT: g11421959, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF237807.1, EVALUE 0.00e+00
US-10-029-386-9711

Query Match 14.3%; Score 215.6; DB 16; Length 530;
Best Local Similarity 98.2%; Pred No. 1.2e-56;
Matches 218; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1286 ACAACGACGATTCGAGTTCCGAGAAAGGAGCTTTGGAGAAACAAATAAGAACCATTTGTGCAG 1345
D 530 ACAACGACGATTCGAGTTCCGAGAAAGGAGCTTTGGAGAAACAAATAAGGCGAGTTGTGCAG 471
QY 1346 GCCTCTCGCAAGAGAAAGAGTGCACACATCTCCACGGCGAAGCGCTCCGTGGAGC 1405
D 470 GCCTCTCGCAAGAGAAAGAGTGCACACATCTCCACGGCGAAGCGCTCCGTGGAGC 411
QY 1406 CGGATCCCAAGTTACCCAGTAGCAGGAGAGGAGCCCACTCAGTCTCCGAGTCCGCGAAG 1465
D 410 CGGATCCCAAGTTACCCAGTAGCAGGAGAGGAGCCCACTCAGTCTCCGAGTCCGCGAAG 351
QY 1466 CTTTCAGAACCTTACCACGAGTGTGTTGCACTCCAGGGTCC 1507
D 350 CTTTCAGAACCTTACCACGAGTGTGTTGCACTCCAGGGTCC 309

RESULT 10
US-10-040-739-520
; Sequence 520, Application US/10040739
; Publication No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John
; Lavallie, Edward
; Racie, Lisa
; Merberg, David
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 520:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 520:
US-10-040-739-520

Query Match          10.4%; Score 156.4; DB 13; Length 469;
Best Local Similarity 99.4%; Pred. No. 5.2e-38;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1350 CTCGCAAGAGAAAGAGAGTGAACACTCCACACGGCGAAGCGCCCTCGTGGAGCGCGA 1409
DB 19 CTAGCAAGAGAAAGAGAGTGAACACTCCACACGGCGAAGCGCCCTCGTGGAGCGCGA 78
QY 1410 TCCCAAGTTACCCAGTAGCAGGAGAGCGCACTCACTGCTCCGAGTGGCGCAAGCTTT 1469
DB 79 TCCCAAGTTACCCAGTAGCAGGAGAGCGCACTCACTGCTCCGAGTGGCGCAAGCTTT 138
QY 1470 CAGAACCTTACCACCACTGCTGTTCTGCACTTCCAGGGTCC 1507
DB 139 CAGAACCTTACCACCACTGCTGTTCTGCACTTCCAGGGTCC 176

RESULT 11
US-10-172-118-1511
; Sequence 1511, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1511
; LENGTH: 6033
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 014643
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1511

Query Match          6.8%; Score 103.2; DB 17; Length 6033;
Best Local Similarity 52.7%; Pred. No. 1.3e-20;
Matches 291; Conservative 3; Mismatches 246; Indels 12; Gaps 3;

QY 697 GAATTACTGTGAGGTATGTGGGAGACATTTAGATCGCTTTTGATGTTGAGATCCAC 756
DB 1057 GAGTTCCTCGGAGGTGTGTGGCCAGGCCTTTCAGCCAGACCTGTTCTCTGAAGCGGCAC 1116
QY 757 ATGAGAACACAAAGATTCTTTTCACTTACGGGTGTAAACATGTGCGGAAGAGATTCAAG 816

; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 520:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 520:
US-10-040-739-520

Query Match          10.4%; Score 156.4; DB 13; Length 469;
Best Local Similarity 99.4%; Pred. No. 5.2e-38;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1350 CTCGCAAGAGAAAGAGAGTGAACACTCCACACGGCGAAGCGCCCTCGTGGAGCGCGA 1409
DB 19 CTAGCAAGAGAAAGAGAGTGAACACTCCACACGGCGAAGCGCCCTCGTGGAGCGCGA 78
QY 1410 TCCCAAGTTACCCAGTAGCAGGAGAGCGCACTCACTGCTCCGAGTGGCGCAAGCTTT 1469
DB 79 TCCCAAGTTACCCAGTAGCAGGAGAGCGCACTCACTGCTCCGAGTGGCGCAAGCTTT 138
QY 1470 CAGAACCTTACCACCACTGCTGTTCTGCACTTCCAGGGTCC 1507
DB 139 CAGAACCTTACCACCACTGCTGTTCTGCACTTCCAGGGTCC 176

RESULT 12
US-10-342-887-1511
; Sequence 1511, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1511
; LENGTH: 6033
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-342-887-1511

Query Match          6.8%; Score 103.2; DB 18; Length 6033;
Best Local Similarity 52.7%; Pred. No. 1.3e-20;
Matches 291; Conservative 3; Mismatches 246; Indels 12; Gaps 3;

QY 697 GAATTACTGTGAGGTATGTGGGAGACATTTAGATCGCTTTTGATGTTGAGATCCAC 756
DB 1057 GAGTTCCTCGGAGGTGTGTGGCCAGGCCTTTCAGCCAGACCTGTTCTCTGAAGCGGCAC 1116
```

QY 757 ATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAAATGTGCGGAAGAAGATTCAAG 816
DB 1117 ATGAAGAAGACACGGGGCTCTTTCGACCACGGCTGCACATCTGGCGCGTAGTTCAAG 1176
QY 817 GAGCCTTGTTTCTTAAATAATCACATGCGGACRCATATGCAAAATCGGGGCCAGAGC 876
DB 1177 GAGCCTTGTTTCTTCAAGAACACATATGAAGGCGACGGCCCCAAGACGCGGACAGAAAC 1236
QY 877 AAATGCGACAGAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCCGTCAGGTGCAC 936
DB 1237 AGGCCCAAGAGTGAGCTGGACCCCA---TCGCCACCATCAACATCTGTCAGGAGGAG 1293
QY 937 GCGGCCGAGAGCATCTCTCTCTCTTGCAGAAATCTGCAATGTTTGGGTCTTCTATTTCOA 996
DB 1294 GTGATCGTCGCGGCTGAGCCTCTACGAGGTCTGCGCAAGTGGCGGGAACCTGTTTACA 1353
QY 997 RATAAAGAAGTCTAATTGAGCACCGCAGGTGCGACACCAACAAAACCTGTTTCGTACC 1056
DB 1354 AACCTGGACAGTTGAACGCCCAATGCAATCCACCGCAGAGTCGAGGC---CAGCCGC 1410
QY 1057 AGCAGCGCGCAGACAGACTCTCCAAAGGAGGAATGCCGTCTCGAGGAGGACTTCTCTG 1116
DB 1411 AGCGGCCCCCGCGAGGAGGGGGCGAGGGCCCTCGACACCAACAGCAGTTCTTCTC 1470
QY 1117 CAGTTGTTCAATTGAGACCAAAATCTCACTGAAACG-----GGGAAGAAGCTGTCT 1170
DB 1471 CAGTGCTGAACTGAGGCGGTGCGCGGCGCGGACTCGTGCCCTGGCACACGAGCTTCTTCTC 1530
QY 1171 AGATCATCCCTCAGTCGATCGGTTCACCACTTCCAGGCTTGGCAKCTGGCTACCAA 1230
DB 1531 CGCGGGTGGCTGAGCTGGACCCGGTCAACAGCTACCAAGGCTGGGAGCTGGCCACGCGG 1590
QY 1231 GGAAGAGTTGCC 1242
DB 1591 GGTAAAGTGGCC 1602

RESULT 13

US-10-074-475-93
; Sequence 93, Application US/10074475
; Publication No. US2003092898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Karra, Kalpana
; APPLICANT: Caferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE REFERENCE: DEX-0313
; CURRENT APPLICATION NUMBER: US/10/074,475
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,292
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 295
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 93
; LENGTH: 8156
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-074-475-93

Query Match 6.8%; Score 103.2; DB 14; Length 8156;

Best Local Similarity 52.7%; Pred. No. 1.6e-20;

Matches 291; Conservative 3; Mismatches 246; Indels 12; Gaps 3;

QY 697 GAATTTAGCTGTGAGTATGTGGGACAGACATTTAGATCGCTTTTGATGTTGAGATCCAC 756

DB 1057 GAGTTCCCGTGCAGGTTGTGGCCAGGCGCTTTCAGCCAGACCTGGTTCTCTGAAGCGGCAC 1116

QY 757 ATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAAATGTGCGGAAGAAGATTCAAG 816
DB 1117 ATGAAGAAGACACGGGGCTCTTTCGACCACGGCTGCACATCTGGCGCGTAGTTCAAG 1176
QY 817 GAGCCTTGTTTCTTAAATAATCACATGCGGACRCATATGCAAAATCGGGGCCAGAGC 876
DB 1177 GAGCCTTGTTTCTTCAAGAACACATATGAAGGCGACGGCCCCAAGACGCGGACAGAAAC 1236
QY 877 AAATGCGACAGAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCCGTCAGGTGCAC 936
DB 1237 AGGCCCAAGAGTGAGCTGGACCCCA---TCGCCACCATCAACATCTGTCAGGAGGAG 1293
QY 937 GCGGCCGAGAGCATCTCTCTCTCTTGCAGAAATCTGCAATGTTTGGGTCTTCTATTTCOA 996
DB 1294 GTGATCGTCGCGGCTGAGCCTCTACGAGGTCTGCGCAAGTGGCGGGAACCTGTTTACA 1353
QY 997 RATAAAGAAGTCTAATTGAGCACCGCAGGTGCGACACCAACAAAACCTGTTTCGTACC 1056
DB 1354 AACCTGGACAGTTGAACGCCCAATGCAATCCACCGCAGAGTCGAGGC---CAGCCGC 1410
QY 1057 AGCAGCGCGCAGACAGACTCTCCAAAGGAGGAATGCCGTCTCGAGGAGGACTTCTCTG 1116
DB 1411 AGCGGCCCCCGCGAGGAGGGGGCGAGGGCCCTCGACACCAACAGCAGTTCTTCTC 1470
QY 1117 CAGTTGTTCAATTGAGACCAAAATCTCACTGAAACG-----GGGAAGAAGCTGTCT 1170
DB 1471 CAGTGCTGAACTGAGGCGGTGCGCGGCGCGGACTCGTGCCCTGGCACACGAGCTTCTTCTC 1530
QY 1171 AGATCATCCCTCAGTCGATCGGTTCACCACTTCCAGGCTTGGCAKCTGGCTACCAA 1230
DB 1531 CGCGGGTGGCTGAGCTGGACCCGGTCAACAGCTACCAAGGCTGGGAGCTGGCCACGCGG 1590
QY 1231 GGAAGAGTTGCC 1242
DB 1591 GGTAAAGTGGCC 1602

RESULT 14

US-10-029-386-23411/c
; Sequence 23411, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23411
; LENGTH: 251
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: EST_HUMAN HIT: W05407.1, EVALUE 1.00e-117
; OTHER INFORMATION: NT HIT: g111421959, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: O75362, EVALUE 1.00e-38
US-10-029-386-23411

Query Match

6.4%; Score 97; DB 16; Length 251;

Best Local Similarity 100.0%; Pred. No. 1.9e-19;

Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1411 CCCAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTC 1470
 |||||
 Db 251 CCCAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTC 192
 |||||

QY 1471 AGAACCTACCCAGCTGGTCTTTCACCTCCAGGGTCC 1507
 |||||
 Db 191 AGAACCTACCCAGCTGGTCTTTCACCTCCAGGGTCC 155
 |||||

RESULT 15

US-10-758-307-110
 ; Sequence 110, Application US/10758307
 ; Publication No. US20040209290A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GENOMIC HEALTH, INC.
 ; APPLICANT: RUSH UNIVERSITY MEDICAL CENTER
 ; APPLICANT: Cobleigh, Melody
 ; APPLICANT: Shak, Steven
 ; APPLICANT: Baker, Joffre
 ; APPLICANT: Cronin, Maureen
 ; TITLE OF INVENTION: GENE EXPRESSION MARKERS FOR BREAST
 ; FILE REFERENCE: 39740/0008 US
 ; CURRENT APPLICATION NUMBER: US/10/758,307
 ; CURRENT FILING DATE: 2004-01-14
 ; PRIOR APPLICATION NUMBER: US 60/440,861
 ; PRIOR FILING DATE: 2003-01-15
 ; NUMBER OF SEQ ID NOS: 440
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 110
 ; LENGTH: 70
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Amplicon
 US-10-758-307-110

Query Match 4.6%; Score 70; DB 20; Length 70;
 Best Local Similarity 100.0%; Pred. No. 2.9e-11;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1419 ACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACCTA 1478
 |||||
 Db 1 ACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACCTA 60
 |||||

QY 1479 CCACCAGCTG 1488
 |||||
 Db 61 CCACCAGCTG 70
 |||||

Search completed: June 29, 2005, 12:01:53
 Job time : 593.497 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 15:06:07 ; Search time 3243.1 Seconds
(without alignments)
17687.639 Million cell updates/sec

Title: US-08-731-499-3
Perfect score: 1507
Sequence: 1 GCAGTGTCTGGATGACT.....GGCTTGCATCCAGGGTCC 1507

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gss1:*
- 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description	
	Score	Match	Length				
1	590.4	39.2	595	4	BF944766	BF944766 PM3-NN117	
2	557.4	37.0	618	5	EX955300	EX955300 DKF2p781N	
3	529.6	35.1	640	4	BM834501	BM834501 K-EST0109	
4	436.6	29.0	571	7	CR753532	CR753532 DKF2p469L	
5	434.4	28.8	895	9	CR094054	CR094054 Forward S	
C	6	434.2	28.8	569	2	AW753609	AW753609 RCL-CT026
7	401	26.6	585	5	BP366185	BP366185 BP366185	
8	394.4	26.2	598	5	BP270262	BP270262 BP270262	
9	389	25.8	583	5	BP229047	BP229047 BP229047	
10	376.2	25.0	636	6	CB457791	CB457791 715904 MA	
C	11	359.2	23.8	594	7	CV409050	CV409050 RCL-HT026
C	12	343.8	22.8	662	6	CB460210	CB460210 720064 MA
13	336	22.3	458	5	BP201868	BP201868 BP201868	
14	302	20.0	352	4	BF963063	BF963063 PM3-NN117	
15	289.6	19.2	768	2	BF237807	BF237807 601842326	
16	285.6	19.0	412	1	AA069107	AA069107 zml0b10.r	
C	17	262.8	17.4	313	2	BF803316	BF803316 RCL-C1011
18	243.4	16.2	307	8	AZ537566	AZ537566 AST-2P596	
C	19	226	15.0	226	1	AA451924	AA451924 aa85e10.s
C	20	224	14.9	304	2	BF803309	BF803309 RCL-C1011
C	21	221.8	14.7	553	1	AJ743841	AJ743841 AJ743841
22	215	14.3	590	6	CD689109	CD689109 EST5631 h	
23	201.6	13.4	567	1	A1552996	A1552996 ve41c05.y	
24	201.4	13.4	885	7	CR432863	CR432863 CR432863	

ALIGNMENTS

RESULT 1
BF944766 595 bp mRNA linear EST 22-JAN-2001
LOCUS PM3-NN1174-181000-001-c08 NN1174 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF944766
ACCESSION BF944766.1 GI:12362041
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 595)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,P.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM3&t2=PM3-NN1174-181000-001-c08&t3=2000-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 585.

FEATURES
Location/Qualifiers
1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN1174"
/note="Organ: nervous normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application

25	178.6	11.9	884	6	CA789554	CA789554 AGENCOURT	
26	160.2	10.6	477	6	CB548508	CB548508 AMGNNUC:M	
27	159	10.6	915	6	CA981247	CA981247 AGENCOURT	
28	140	9.3	647	1	AL595898	AL595898 AL595898	
29	138.4	9.2	818	9	CR127300	CR127300 Forward s	
30	137.4	9.1	376	1	AA646529	AA646529 ve41c05.f	
31	132.8	8.8	735	5	BP675362	BP675362 BP675362	
c	32	130.8	8.7	505	1	A1467614	A1467614 ve41c05.x
33	130.2	8.6	603	4	BJ031636	BJ031636 BJ031636	
34	122.4	8.1	966	5	BU912448	BU912448 AGENCOURT	
35	118.8	7.9	692	2	BB617224	BB617224 BB617224	
36	117.2	7.8	2331	3	AK030249	AK030249 Mus muscu	
37	116.2	7.7	672	5	EX854767	EX854767 EX854767	
38	114.8	7.6	342	5	BY206008	BY206008 BY206008	
39	114.8	7.6	361	5	BY207082	BY207082 BY207082	
40	114.8	7.6	377	5	BY200588	BY200588 BY200588	
41	111	7.4	338	4	BG083784	BG083784 H3092A10-	
42	108.2	7.2	133	2	BE706424	BE706424 RCL-HT026	
43	104.4	6.9	793	6	CA752347	CA752347 UI-M-FOO-	
44	104	6.9	805	6	CB522302	CB522302 UI-M-GHO-	
45	96.2	6.4	295	1	AJ743842	AJ743842 AJ743842	

No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

```
Query Match      39.2%; Score 590.4; DB 4; Length 595;
Best Local Similarity 99.5%; Pred. No. 3.3e-164;
Matches 591; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 615 GGTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGACAGAACCTCCCAA 674
Db 1 GGTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGACAGAACCTCCCAA 60

QY 675 GGAAGAATTCGAGGAAATGAATTTAGCTGTGAGGTATGTGGCGACACATTTAGAGT 734
Db 61 GGAAGAATTCGAGGAAATGAATTTAGCTGTGAGGTATGTGGCGACACATTTAGAGT 120

QY 735 CGCTTTTGATGTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGGTGTA 794
Db 121 CGCTTTTGATGTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGGTGTA 180

QY 795 CATGTGCGGAAGAAGATCAAGGACCTTGGTTCTTAAATATCATATCGGACACATAA 854
Db 181 CATGTGCGGAAGAAGATCAAGGACCTTGGTTCTTAAATATCATATCGGACACATAA 240

QY 855 TGGCAATCGGGGGCCAGNAGAACTGACGAAAGCTTGGAGATGTCCAGCAACCAT 914
Db 241 TGGCAATCGGGGGCCAGNAGAACTGACGAAAGCTTGGAGATGTCCAGCAACCAT 300

QY 915 CAACGAGGTGCTCAGGTGACGCGCGGAGAGCATCTCTCTCTTGCAAAATCTGCAT 974
Db 301 CAACGAGGTGCTCAGGTGACGCGCGGAGAGCATCTCTCTCTTGCAAAATCTGCAT 360

QY 975 GGTGTGCGCTTCTTATTTCCAAATTAAGAAAGTCTAATTTAGACCCGCAAGGTGCAC 1034
Db 361 GGTGTGCGCTTCTTATTTCCAAATTAAGAAAGTCTAATTTAGACCCGCAAGGTGCAC 420

QY 1035 CAAAAAACTGTTTCGGTACACGAGCGCGAGAGCATCTTCCACAGGAGGATGCC 1094
Db 421 CAAAAAACTGTTTCGGTACACGAGCGCGAGAGCATCTTCCACAGGAGGATGCC 480

QY 1095 GTCCCTCGAGGAGGACTTCTCAGTGTGTTCAATTTGAGACCAAAATCTCACCCCTGA 1154
Db 481 GTCCCTCGAGGAGGACTTCTCAGTGTGTTCAATTTGAGACCAAAATCTCACCCCTGA 540

QY 1155 GGGGAAGAAGCTGTGAGATGATCCCTCAGTCCGATCGTTCACACCTTCCA 1208
Db 541 GGGGAAGAAGCTGTGAGATGATCCCTCAGTCCGATCGTTCACACCTTCCA 594
```

RESULT 2

EX955300
LOCUS
DEFINITION DKFZp781N0270 r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
KEYWORDS DKFZp781N0270 5', mRNA sequence.

ACCESSION EX955300

VERSION EX955300.1 GI:43437635

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 618)

Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,

Fobo,G., Han,M. and Wiemann,S.

EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)

Unpublished (2003)

Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@kfz- heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.

No sl sequence available.

This clone (DKFZp781N0270) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..618

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp781N0270"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="781 (synonym: hlcc4)"

/note="Vector: pSport1_SfiI; Site_1: SfiIA; Site_2: SfiIB;

cDNA-collection"

ORIGIN

```
Query Match      37.0%; Score 557.4; DB 5; Length 618;
Best Local Similarity 95.8%; Pred. No. 2.3e-154;
Matches 592; Conservative 1; Mismatches 23; Indels 2; Gaps 2;

QY 377 GGCCAAAGATGATTTGGCAGCTCTCTTGCAGTCCGATGGAGATGGA-GATGCTTTGTCAA 435
Db 2 GGCCAAAGATGATTTGGCAGCTCTCTTGCAGTCCGATGGAGATGGAAGATGCTTTGTCAA 61

QY 436 TGAAGAGGCCCTCTGTTCTCAATTCGAGCTACACAAAGAAAAAATGTCAATCCGAATC 495
Db 62 TGAAGAGGCCCTCTGTTCTCAATTCGAGCTACACAAAGAAAAAATGTCAATCCGAATC 120

QY 496 GAGGGGAATATGCCCTTGATTCATGTCCTGTCAGCCAGACCTTCACATTCAGAAGC 555
Db 121 GAGGGGAATATGCCCTTGATTCATGTCCTGTCAGCCAGACCTTCACATTCAGAAGC 180

QY 556 CTTAATAAACATGTCTTAATGCAACACCGGCTACCCCTCTGTGAACACGAGCTTCTTCGG 615
Db 181 CTTAATAAACATGTCTTAATGCAACACCGGCTACCCCTCTGTGAACACGAGCTTCTTCGG 240

QY 616 GTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGAACCTCCCAAG 675
Db 241 GTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGAACCTCCCAAG 300

QY 676 GAAAGAAATTCAGAGAAATGAATTTAGCTGTGAGTATGTGGCAGACATTTAGAGTC 735
Db 301 GAAAGAAATTCAGAGAAATGAATTTAGCTGTGAGTATGTGGCAGACATTTAGAGTC 360

QY 736 GCTTTTGATTTGAGATCCACATGAGAACACACAAAGATTTCTTTCACTTACGGGTGTAAC 795
Db 361 GCTTTTGATTTGAGATCCACATGAGAACACACAAAGATTTCTTTCACTTACGGGTGTAAC 420

QY 796 ATGTGCGGAAGAAGATTCAGAGAGCTTGGTTTCTTAAAAATTCATGCGGACACATAT 855
Db 421 ATGTGCGGAAGAAGATTCAGAGAGCTTGGTTTCTTAAAAATTCATGCGGACACATAT 480

QY 856 GGCAAAATCGGGGGCCAGAACAACTGCAGCAAGCTTGGAGAGTAGTCCACACAGCATC 915
Db 481 GGCAAAATCGGGGGCCAGAACAACTGCAGCAAGCTTGGAGAGTAGTCCACACAGCATC 540

QY 916 AACGAGGTGCTCCAGGTGCAGCGGCCGAGAGACATCTCTCTCTTCCCTTAAAAATCTGCATG 975
Db 541 AACGAGGTGCTCCAGGTGCAGCGGCCGAGAGACATCTCTCTCTTCCCTTAAAAATCTGCATG 600

QY 976 GTTTGTGGCTTCCCTATTT 993
Db 601 GTTTGTGGCTTCCCTATTT 618
```

RESULT 3

BM834501

LOCUS

DEFINITION

640 bp mRNA linear EST 06-MAR-2002
K-EST0109502 S11SNUI Homo sapiens cDNA clone S11SNUI-63-D04 5',

```

mRNA sequence.
BM834501      1  GI:19190910
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 640)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 63 row: D column: 04
High quality sequence stop: 640.
Location/Qualifiers
1..640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S11SNUI-63-D04"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/lab_host="SNU-1"
/clone_lib="S11SNUI"
/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested pME18S-FL3 vector. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

FEATURES
source
1..571
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469L046"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469"
/note="Vector: pSport1_Sfi; Site 1: SfiI; Site 2: SfiI"

ORIGIN
Query Match 35.1%; Score 529.6; DB 4; Length 640;
Best Local Similarity 95.2%; Pred. No. 4.6e-146;
Matches 610; Conservative 0; Mismatches 25; Indels 6; Gaps 6;

QY 67 TCCTTTGCTCATTTCAATTTGAATTCGAGGCAATGGGAACCTGTATGCCCTTTGTTGG 126
DB 1 TCCTTTGCTCATTTCAATTTGAATTCGAGGCAATGGGAACCTGTATGCCCTTTGTTGG 60

QY 127 GAAAGAACAGGTGACACCATCACTAGCTTCTCTAAAGTTGGAAGTTAGAGACTAT 186
DB 61 GAAAGAACAGGTGACACCATCACTAGCTTCTCTAAAGTTGGAAGTTAGAGACTAT 120

QY 187 ACACITTTCTTTTGAACCTTTTATAATAATTTGCTCTGG-TTTTGGACCCGAGCTGT 245
DB 121 ACACITTTCTTTTGAACCTTTTATAATAATTTGCTCTGGTTTGGACCCGAGCTGT 180

QY 246 TAGA-GGGTGAGTGACAGGTCTTAC-AGTGCCCTTAATCCAACCTCAGAAAATGGCCCAAC 303
DB 181 TAGAGGGGTGAGTGACAGGTCTTACAGTGGGCTTTATTCCAACTCAGAAAATGGCCCAAC 240

QY 304 GGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTT 363

```

```

Db 241 GGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTT 300
QY 364 AATGTACATGGAT-GGCCCAAGAGTATTGGGAGCTCTCTTGGCAGTCCGATCGAGATGGA 422
Db 301 AATGTACATGGATGGCCCAAGAGTATTGGGAGCTCTCTTGGCAGTCCGATCGAGATGGA 360
QY 423 -GATGCCCTTGTCAATGAAGAGGCCCNCTGTGTTCATTTCCGAGCTACACAAAGAAAAAAA 481
Db 361 GGATGCCCTTGTCAATGAAGAGGCCCGCTGTGTTCATTTCCGAGCTACACAAAGAAAAAAA 420
QY 482 TGTCAATCGATCGAGGGGATATGCCCTTGGATGTTCGTCAGCCAGACCTTCA 541
Db 421 TGTTC-ATCCAAATCGAGGGGTATATGCCCTTGGATGTTCGTCAGCCAGACCTTCA 479
QY 542 CACATTGAGAAGACCTTAATAAAACATGTCTTAATGCAACACCGGCTACCCCTCTGTGAAC 601
Db 480 CACATTGAGAAGACCTTAATAAAACATGTCTTAATGCAACACCGGCTACCCCTCTGTGAAC 539
QY 602 CAGCAGTTCTTCGGGTTCAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTCCGAA 661
Db 540 CAGCAGTTCTTCGGGTTCAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTCCGAA 599
QY 662 CAGACCTTCCCAAGCAAGAAAGATTCCAGGAAATGAATTT 702
Db 600 CAGACCTTCCCAAGCAAGAAAGATTCCAGGAAATGAATTT 640

RESULT 4
CR753532      571 bp  mRNA  linear  EST 01-SEP-2004
LOCUS DKFZp469L046_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
DEFINITION DKFZp469L046 5', mRNA sequence.
ACCESSION CR753532
VERSION CR753532.1 GI:51845942
KEYWORDS EST.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
REFERENCE 1 (bases 1 to 571)
AUTHORS Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
TITLE Pongo pygmaeus mRNA (Wambutt,R., Heubner,D., Mewes,H.W., et al.)
JOURNAL Unpublished (2004)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert. Clone from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email s.wiemann@dkfz-heidelberg.de; sequenced by Agowa
(Berlin/Germany) within the cDNA sequencing consortium of the
German Genome Project. This clone (DKFZp469L046) is available at
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469L046
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
Location/Qualifiers
1..571
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469L046"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469"
/note="Vector: pSport1_Sfi; Site 1: SfiI; Site 2: SfiI"

ORIGIN
Query Match 29.0%; Score 436.6; DB 7; Length 571;
Best Local Similarity 92.0%; Pred. No. 2.5e-118;

```


Matches	526;	Conservative	0;	Mismatches	40;	Indels	6;	Gaps	6;
QY	56	AAGAGCACTAATGCTTTGTGCTGATTTCATATTTGAATCGAGGCATTTGGAAACCCCTGTATG	115						
Db	1	AAGAGAACTAATGCTTTGTGCTGATTTCATATTTGAATCAAGGCATTTGGAGCCCTGTATG	60						
QY	116	CCTCTGTTGTGAAAGAACCCAGTGACACCATCACCTGAGCTTCTCTAAAGTTTCGAAGAAGT	175						
Db	61	CGTGTTTGTGGAAGAACAACAGTGACACCACTGAGCTTCTCTAAAGTTTCGAAGAAGT	120						
QY	176	TAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAATTTTCTCTCTGGTTTTTGGAA-	234						
Db	121	TAGAGGACTATACGCTTTCTTTTGAACCTTTTATAATAATTTTCTCTCTGGTTTTTGGAA	180						
QY	235	CCAGGACTGTTAGA-GGGTGAGTGACAGGCTTTAC-AGTGGCCCTTAATCCAACTCCAGA	292						
Db	181	CCAGGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCCTTATTTCCAACCTCCAGA	240						
QY	293	AATTGCCCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCAACT	352						
Db	241	AATTCCCCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCAACT	300						
QY	353	CAATCCCTCTTAATGTACATGAT-GGCCAAGAGTGATTGGCAGCTCTCTTGCCAGTCCG	411						
Db	301	CAGTCCCTCTTAATGTACATGATGGCCAGAGTGATTGGCAGCTCTCTTGCCAGTCCAG	360						
QY	412	ATGGAGATGGA-GATGCCCTTGTCAATGAAGGGCCNCCTGTGTCAATTCGAGGTACAC	470						
Db	361	ATGGAGATGGAGGAGCGCCATGTCAATGAAGGGACCGCTGTGTTCATTCGAGCTACG	420						
QY	471	AAAGAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCAG	530						
Db	421	CAAGAAAAAATGTCT-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAG	479						
QY	531	CCAGACCTTTCACATTCAGAGACCTTAATAAATCATGTCTTAATGCAACACCGGCCTAC	590						
Db	480	CCAGACCTTTCACATTCAGAGACCTTAATAAATCATGTCTTAATGCAACACCGGCCTAC	539						
QY	591	CCTCTGTGAACAGCAGTCTTCGGGTTGAAG	622						
Db	540	CCTCTGTGAACAGCAGTCTTCGGGTTGAAG	571						

RESULT 5
 CR094054
 LOCUS
 DEFINITION
 Forward strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN19116, genomic survey sequence.
 CR094054
 ACCESSION
 CR094054.1 GI:49827883
 VERSION
 GSS; genome survey sequence; MICR.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 895)
 Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
 Direct Submission
 Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 TITLE
 JOURNAL
 CB10 15A, UK. http://www.sanger.ac.uk/MICR
 FEATURES
 Location/Qualifiers
 1..895
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHPN19116"
 /clone_lib="MHPN"

ORIGIN

Query Match 28.8%; Score 434.4; DB 9; Length 895;
 Best Local Similarity 69.0%; Pred. No. 1.3e-117;

Matches	620;	Conservative	2;	Mismatches	270;	Indels	6;	Gaps	2;
QY	332	GTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGCCAGAGTGATG	391						
Db	1	GTGGTTGGCAGCATGCCGACCCAGTCCCTCTCGTGTCTATGACCGGCGGAAGTCTCTCA	60						
QY	392	GCAGCTCTCTTGGCCAGTCCGATGGAGATGGAGTTCCTTGTCAATGAAGGCCCCNCCTGT	451						
Db	61	GCAGCTCTCTAGGTCCCGATGGAGTGGATGTCTGTGCCCATAAAGGCCCGGTGG	120						
QY	452	TGTCAATTTCCGAGCTTACAAAGAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCT	511						
Db	121	CAGTCCCTCTCCGAGCTCTCAGGAGAAAGAGCATGGCCGTGGCAGAGGCCACATGCCCC	180						
QY	512	TGGATTGATGTTCTGAGCCAGACCTTTCACATTCAGAAGACCTTAATAAATCATGTCT	571						
Db	181	TGGATTGATGTTCTGAGCCAGGCTCTCTCTCAGGGCGGAGATCTCTAGTCAGCACGTGC	240						
QY	572	TAATGCAACACCGGCTTACCCTCTGTGAAACACGACGATTTCTTCGGGTTCGAAGCAGAGTATC	631						
Db	241	TGCTGCAGCACCGGCCACCTCTGCGAGCCAGCTGTCTCGGTGTGGAGGCCGAGTACC	300						
QY	632	TCAGTCCCTTTGATAAAAGTCAAGTGCAGAACAGAACTCTCCAAAGAAAAAAGTAATCCAAAG	691						
Db	301	TAAGTCCCTTTGATAAAAGTCTGAGAGCCAAACAGAGCCAGCATTTGGAGAAGAGTGGCGAAG	360						
QY	692	A---AAATGAATTTAGCTGTGAGTATGTGGCAGACATTTAGAGTCGCTTTTGTATGTTG	748						
Db	361	ACCCGAGGAGTGTGAGCTGTGATGTGTGGCGACATTTCCACGTGCTTTTGTATGTTG	420						
QY	749	AGATCCACATGAGAAACACAAAAAGATTTCTTTCACCTTACCGGGTGTAAACATGTGCGGAAGAA	808						
Db	421	AGAGCCATGAGAGAGCATTAAGACTCTCTTCACTATGGGTGACGATGTGCGGGAGGA	480						
QY	809	GATTCAAGAGCCCTTGGTTTCTTAAAAATCATATCGGACRCATATATGCGAAATCGGGG	868						
Db	481	GATTCAAGAGCCCTTGGTTTCTTAAAGAACACATGCGGACACACAATGGCAAGTCTGGCA	540						
QY	869	CCAGAGCAAACTGACGCAAGGCTTGGAGATGTCTCCAGCAACGATCAACGAGTCTGCC	928						
Db	541	CCAGAGCAAACTGACGCAAGGCTTGGAGATGTCTCCAGCAACGATCAACGAGTCTGCC	977						
QY	929	AGGTGACGCGCGCGAGAGCATCTCTCTCTTGCATAATCTGCATGTGTGTGCTTCC	988						
Db	598	AGCGCAGCCCTTGGGAGCATCTCCAGCCCTACAGATCTGCATGTCTGCGGCTTCC	657						
QY	989	TATTTCCAAATAAAGAAAGTCTAATTTGAGCACCGCAAGGTGCACACCAAAAAAATCTGTT	1048						
Db	658	TCTTCCCAATAAAGCAGAGCCCTCATTTGAGCACAGCAAGGTTTACCGCCAAAGAAATCTGCC	717						
QY	1049	TCGGTACCGAGCGCGCAGACAGACTCTCCACAGGAGGATGCCGTCTCGAGGGAGG	1108						
Db	718	CCAGTGCCAGCAACGTTTGCCTCTGATGATCCACCGAGAGAAACCCACGTCCTCCGAGGAG	777						
QY	1109	ACTTCTCTCAGTGTGTTCAACTTTGAGACCAAAAATCTCACCTCTGAAACGGGGGAAGACCTG	1168						
Db	778	AGTTGCTGCACTTTTGAATTTGAGACCCAGATCAATGCGANGTAGTACAGTGAAGCCCA	837						
QY	1169	TCAGATGATCCCTCAGCTCGATCGTTTACACACCTTCCAGGCTTGGCAKCTGGGTAC	1226						
Db	838	TGACCTGCATACCTCAGCTTGACCCGTTTACACACCTTACCGAGCATGGCAGTTGGCTAC	895						

RESULT 6
 AW753609/c
 LOCUS
 DEFINITION
 RC1-CT0268-060100-013-h08 CT0268 Homo sapiens cDNA, mRNA sequence.
 ACCESSION
 AW753609
 VERSION
 AW753609.1 GI:7668541
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 569)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
TITLE
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
JOURNAL
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
 20202663
PUBMED
 10737800
COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC1&t2=RC1-CT0268-
 060100-013-h08&t3=2000-01-06&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 49
 High quality sequence stop: 67.
FEATURES
 source
 1..569
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="CT0268"
 /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
 SmaI; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No.
 196,716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
ORIGIN
 Query Match 28.8%; Score 434.2; DB 2; Length 569;
 Best Local Similarity 89.5%; Pred. No. 1.3e-117;
 Matches 510; Conservative 1; Mismatches 55; Indels 4; Gaps 4;
 QY 430 TGTCAATGAAGGGCCNCCTGTGTGTCATTCGGAGCTACACAAAGAAAAATGTCAATC 489
 DB 569 TGTCCATGAAGGGACCGCTGTGTTCATTCGGAGCGACACAGAATAAGATGTC-ATC 511
 QY 490 CGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCA-CCGACACCTTCACACATTC 548
 DB 510 CCACCTCGAGGGGTATATGCCCTTGGATGCATGTTCTGCATGCCAGACCTGCACACATTC 451
 QY 549 AGAAGACCTTAATAACATGTCTTAATGCAACACCGGCCCTAGCC-TCTGTGAACACAGCAG 607
 DB 450 AGAAGACCTGAATCCACATGTCTGATGCGACACCGGCCCTACCCCTCTGTGAACACAGCAG 391
 QY 608 TTCTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGGCAACAGAAC 667
 DB 390 TTCTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGGCAACAGAAC 331
 QY 668 CTCGCCAGAAAGAAATTCAGAGAAATGAATTTAGCTGTGAGGTATGTGGGCGACAT 727
 DB 330 CTCGCCAGAAAGAAATTCAGAGAAATGAATTTAGCTGTGAGGTATGTGGGCGACAT 271
 QY 728 TTAGAGTCCCTTTTGTGATTTGAGATCCACATGAGAACACACAAAGATCTTTCACCTACG 787
 DB 270 TTAGAGTCCCTTTTGTGATTTGAGATCCACATGAGAACACACAAAGATTCGCTACCTGACG 211
 QY 788 GGTGTAACATGTCGGAAGAAGATTCAAGGACGCTTGGTTCTTTAAANAATCAATGCGGA 847

DB 210 GGTGTGACATGTGCGGTAGAAGATTTAAGGAGCCTTGGTTCTTAATAATCACATCGGA 151
 QY 848 CRCATAATGGCAAAATCGGGGGCCAGAACTGCAGCAAGCCTTGGAGAGTAGTCCAG 907
 DB 150 CACATAATGGCAAAATCGGGGGCCAGAACTGCAGCAAGCCTTGGAGAGTAGTCCAG 91
 QY 908 CAACGATCAACAGAGTCTGTCAGGTCGACGCGCGGAGAGCATCTCTCTCC-TTGCAAA 966
 DB 90 CCACGATCAACAGAGTCTGTCAGGTCGACGCGCGGAGAGCATCTCTCTCTCGTTACCA 31
 QY 967 ATCTGCATGTTTGGCTTCTCTATTCCA 996
 DB 30 ATCTGCATGTTTCCGTTCTCTAGTTCCA 1
RESULT 7
 BP366185
LOCUS
 DEFINITION
 sapiens cDNA clone TIR09337, mRNA sequence.
ACCESSION
 BP366185
VERSION
 BP366185.1 GI:52296435
KEYWORDS
 EST.
SOURCE
 Homo sapiens (human)
ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 1 (bases 1 to 585)
 Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
 Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE
 Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
JOURNAL
 Genome Res. 14 (9), 1711-1718 (2004)
COMMENT
 Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yszukui@ims.u-tokyo.ac.jp.
FEATURES
 source
 1..585
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="TIR09337"
 /tissue_type="lung"
 /cell_type="fibroblast"
 /cell_lines="TIG"
 /dev_stage="fetal"
 /clone_lib="Sugano cDNA library, fetal lung fibroblast
 TIG"
ORIGIN
 Query Match 26.6%; Score 401; DB 5; Length 585;
 Best Local Similarity 93.8%; Pred. No. 1e-107;
 Matches 482; Conservative 0; Mismatches 26; Indels 6; Gaps 6;
 QY 2 CAGGTGTCTGGATGACCTTCTTGCTCAATTGAACACCTCATTCATGAGACAAGAGC 61
 DB 73 CAGGTGTCTGGATGACCTTCTTGCTCAATTGAACACCTCATTCATGAGACAAGAGC 132
 QY 62 ACTAATGCTTTGTGCTGATTTCATATTGAATCGAGGCATTTGGAAACCCCTGTATGCTTGT 121
 DB 133 ACTAATGCTTTGTGCTGATTTCATATTGAATCGAGGCATTTGGAAACCCCTGTATGCTTGT 192
 QY 122 TTGTGGAAGAACACAGTGACACCATCATGAGCTCTTAAAGTTTGAAGAGTTAGAGG 181
 DB 193 TTGTGGAAGAACACAGTGACACCATCATGAGCTCTTAAAGTTTGAAGAGTTAGAGG 252
 QY 182 ACTATACACTTTCTTTTGAACCTTTTATAATAATATTCCTCTGG-TTTTGGAAACCCAGG 240
 DB 253 ACTATACACTTTCTTTTGAACCTTTTATAATAATATTCCTCTGGTTTGTGGAACCCAGG 312


```

286 ACTATACACTTTCTTTTGAACCTTTTATATAATAATATTTGCTCTGGTTTGTGAAACCCAGG 345
241 ACTGTTTAGA-GGGTGAAGTACAGAGTCTTAC-AGTGGCCTTAATCCAATCCAGAAATTCG 298
346 GCTGTTAGAGGGTGAAGTACAGAGTCTTACAGTGGCCTTATCCAACTCCAGAAATTCG 405
299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCAACTCAATCC 358
406 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCAACTCAATCC 465
359 CTCCTTAATGTACATGGAT-GGCCAAGAGTGATGGCAGCTCTCTTCCAGTCCGATGGAG 417
466 CTCCTTAATGTACATGGATGGCCAGAGTGATGGCAGCTCTCTTGGCAGTCCGATGGAG 525
418 ATGGA-GATGCCCTTGTCAATGAAGGGCCNCTGTG-TCAATTCGGAGCTACACAAA 473
526 ATGGAGGATGCTTGTCAATGAAGGGCCGCTGTGTTCATTCGAGCTACACAAA 583

RESULT 10
CB457791
LOCUS CB457791 636 bp mRNA linear EST 26-MAR-2003
DEFINITION 715904 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION CB457791
VERSION CB457791.1 GI:29264173
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,
AUTHORS Wray,J.E. and Keefe,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross match v0.990329.
Plate: FQ8073 Row: A Column: 4
Seq primer: GTAATACGACTCATATAGG.
FEATURES
source
location/Qualifiers
1..636
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."
ORIGIN
Query Match 25.0%; Score 376.2; DB 6; Length 636;
Best Local Similarity 76.4%; Pred. No. 2.7e-100;
Matches 488; Conservative 1; Mismatches 144; Indels 6; Gaps 2;
496 GAGGGGAATATGCCCTTGATTCAGTTCCTGACGCCAGACTTCACACATTTCAGAAAC 555
1 GAGGGGTGCATGCCCTTGACCTGCATGTCTTCAGTCAGAGCTTCACGCGCTCCGAGAC 60
556 CTTAATAACATGCTTTAATGCAACACCGCCCTACCTCTGTGAACACGAGCTTTTCGG 615
61 CTCAGCAAGCAGCTCTCTCTGACGACCGGCCACCTCTGTGAGCCTGCGCTCCTCGCG 120

```

```

616 GTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCTAGTGCAGACAGAACTCTCCCAAG 675
121 GTGAGGCGGAGTATCTCAGCCCTCTGGATAAAGTCCAGGTGCGCAGCGAGCTCTCCCAAG 180
676 GAAAAGAAATTCAGAGAAATGAA---TTTACTGTAGGTATGTGGGAGACAGATTAGA 732
181 GACAAAGAGCTGCAGGAGAACGAGAGCTGAGCTGTGAAGTGTGCGGGCAGACGTTCCGC 240
733 GTGCGTTTGTGTTGAGATCCACATGAGAACACACAAAGATTCTTTCACTTACGGGTGT 792
241 GTGGCTTTTCGACTCGAGATCCACATGAAGAAACACAAAGACTCGTTCACTGACGGGTGT 300
793 AACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAATAATCACATGCGGACRAT 852
301 CACATGTGCGGCGCGGTTCGAAGGAGCCTGGTTTCTTGAAGAAATCACATGCGGACGAC 360
853 AATGCCAATCGGGGCGCAGAGCAAACTGACAGCAAGGCTTGAGAGTAGTTCAGCAAG 912
361 ACGGCAAGTCCGGGGCGAAGAGCATGCTGACGCCGCCCTGGAG---AGCCCCGCCACC 417
913 ATCAACGAGGTGCTCCAGGTGCAAGCGCGCAGAGACATCTCTCTCTCTGCAAAATCTGC 972
418 ATCAACGAGGTGCTCCAGAGCCCGCGCGAGGCGCTCTCTCGCGGTATAGATATGC 477
973 ATGGTTTGTGCTCTCTATTTCCAAATAAAGAAAGTCTAAATTGAGCACCGCAAGGTGCAC 1032
478 ATGGTCTCGGCTTCTGTTTCCAAATAAAGAAAGTCTCATCGATCACCAGGATGCAC 537
1033 ACCAAAAAATCGCTTTCGTTACAGAGCGCGCAGACAGACTCTCCACAGAGGAATG 1092
538 ACCAAAGAGTGTCTTCCGGCGCCGCGCAGAGTCGAAGAGCCCAAGAGGAGCATG 597
1093 CGTCTCTCAGGAGGAGTCTTCTGAGTTGTTCAACTTG 1131
598 CTGTCGCCGGGGCGGAGCTGCTGCGAGTTCCTGAATCTG 636

RESULT 11
CV409050/c
LOCUS RCI-HR0268-090500-017-c10 HT0268 Homo sapiens cDNA, mRNA sequence.
DEFINITION CV409050
ACCESSION CV409050.1 GI:52804553
VERSION CV409050.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 594)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc.Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
MEDLINE 10737800
PUBMED
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. http://www.ludwig.org.br.
location/Qualifiers
1..594
source

```

/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /dev_stage="Adult" /clone_lib="HT0268" /note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."		ORIGIN						
Query Match	23.8%; Score 359.2; DB 7; Length 594;							
Best Local Similarity	84.9%; Pred. No. 3.1e-95;							
Matches 485; Conservative	1; Mismatches 70; Indels 15; Gaps 7;							
QY	423 GATGCTTTGTCAATCAAAAGGGCCNCTGTTG-TCAATTCGGAGCTACACAAAGAAAAA 481							
DB	581 GATGCTTTGGCAATGAAGGACCCCTGTTTCATTCGGAGCTACACAGGAAGAAT 522							
QY	482 TGTCAATCCGAATCAGAGGGGAATATGCCCTTGG-----ATTGCATGTTCTGCAGCCAGACC 537							
DB	521 GGTCAATCCAAATCTCAGGGTGATATGCCCTTGGGATTTTCATGGTTGCTGCAGCCAGACC 462							
QY	538 TTCACACAT-----TCAGAAGACCTTAATAACATGCTTAATGCAACACCGGC-CTACCC 592							
DB	461 TTCACACATGTCAGAGAGACCTTGAAGATACATGCTTAATGCAACACCGGCACCTACCC 402							
QY	593 TCTGTGAACACGACGCTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTC 652							
DB	401 TCTGTGAACACGACGCTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTC 342							
QY	653 AAGT-GCGAACAGAACCTCCCAAGAAAGAAATTCGAAGGAAAATGAATTTAGCTGTGAG 711							
DB	341 AAGTGGCCAAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAAATGAATTTAGCTGTGAG 282							
QY	712 GTATGTGGGCAGACATTTAGAGTCGCTTTGATGTTGATGATCCACATGAGAACACAAA 771							
DB	281 GTATGTGTCAGACATTTAGAGTCGCTTTGATGTTGATGATCCACATGAGAACACAAA 222							
QY	772 GATTTCTTTCACATTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTT 831							
DB	221 GAGTCTTTCACATTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTT 162							
QY	832 AAAAATCATATCGGACACATATATGCGCAAAATCGGGGGCCAG--AAGCAAACTGCAGCAAG 889							
DB	161 ATAAATCAACAGCGGACACATAATGGCTATGCGGGGGCCAGATAGCTATCTGTCAAGCAAG 102							
QY	890 GCTTGGAGATGATCCAGCAAGATCAACAGGTC--GTCCAGGTGACCGGGCCGAGAG 947							
DB	101 GCGTGGAGATGATGTCACAGCATCAACAGATTCTGTCCAGTGTGCACGGGGCCGAGAG 42							
QY	948 CATCTCTCTCTCTTGCAGAAATCTGCATGTTT 978							
DB	41 CATCTCTCTCTCTGACATGATCTGCATGTTT 11							
RESULT 12								
CB460210/c								
LOCUS	CB460210 662 bp mRNA linear EST 26-MAR-2003							
DEFINITION	720064 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.							
ACCESSION	CB460210							
VERSION	CB460210.1 GI:29266594							
KEYWORDS	EST.							
SOURCE	Bos taurus (cow)							
ORGANISM	Bos taurus							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;								
Bovinae; Bos.								
REFERENCE	1 (bases 1 to 662)							
AUTHORS	Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,							
TITLE								
JOURNAL								
COMMENT								
Wray,J.E. and Keele,J.W. A second set of bovine ESTs from pooled-tissue normalized libraries Unpublished (2003) Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smithem@mail.marc.usda.gov Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross_match v0.990329. Plate: FOY8073 row: A column: 4 Seq primer: TAGAAGGCACAGTCGAGG. FEATURES								
source								
1..662								
/organism="Bos taurus"								
/mol_type="mRNA"								
/db_xref="taxon:9913"								
/tissue_type="pooled"								
/lab_host="DH10B"								
/clone_lib="MARC 6BOV"								
/note="vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI; Library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."								
ORIGIN								
Query Match	22.8%; Score 343.8; DB 6; Length 662;							
Best Local Similarity	75.3%; Pred. No. 1.3e-90;							
Matches 466; Conservative	3; Mismatches 140; Indels 10; Gaps 3;							
QY	703 AGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGA-TCCACATGAG 761							
DB	661 AGCTGTGAAGTGTGCGGGCAGACGTTCCGCGTGGCTTCGACGTGCAGATTCACATGAA 602							
QY	762 AACACAAAGAATCTTTTCACTTACGGGTGTAAATCATGTGCGGAAGAAGATTCAAGGAGCC 821							
DB	601 GAACACAAAGNACTCGTTTACGTACGGGTGTACATGTGCGGCGCGGCTTCAAGGAGCC 542							
QY	822 TTGGTTTCTTAAAAATCATATGCGGACACATAATGGCAAAATCGGGGGCCAGAGCAAACT 881							
DB	541 CTGGTTTCTGAAGAATCATATGCGGACGACACACGCGCAAGTCGGGGGGCGAAGAGCAGGT 482							
QY	882 GCAGCAAGCTTGGAGTAGTCCAGCAACCATCAACAGGTCTGTCAGGTGCACGCGGC 941							
DB	481 GGAGCCCGCTTGGAG---AGCCCGGCGACCATCAACAGGTGTGTCAGGAGCCCGCGGC 425							
QY	942 CGAGACATCTCCTCTCTTGCAAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAA 1001							
DB	424 CGGGGCGTCTCTCTCGCGGTATAGATATGATGCTCTGCGCTTCTTGTGTTTCCAAATAA 365							
QY	1002 AGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAATCTGTTTCGGTACAGCAG 1061							
DB	364 AGAAAGTCTCATCGATCACCCGGAAGATGCACCAAAAGAAGCTGCTTCGGGCGCCGCTGG 305							
QY	1062 CGCGCAGACAGACTCTCCACAGGAGGAATGCCGCTCTCGAGGGAGGACTTCTCTGCAGTT 1121							
DB	304 CCCGAGATCGAAGCCCCACAGGAAGGATGCTGTGCGCGGGGGCGGAGCTGCTGCAAGTT 245							
QY	1122 GTTCAACTTTGAGACCAAAATCTCACCTGAAACGGGGGAAGAGCCTGTGCAGATGCATCCC 1181							
DB	244 CTTGAATCTGAGACCCACGCTCCCGCGCGAGATGCCAAGAAGCCGCGCAAGTGATACC 185							
QY	1182 TCAGCTCGATCGTTTACCACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAWAGTTGC 1241							
DB	184 TCAGCTGAGCCCGTTTCAACCTTACAGGCTTGGCAGCTGGCCACCAAGGGCAAGGTGGC 125							
QY	1242 CATTTGCCAAGAAGTGAAGG-----AATTGGGGCAAGAAGGAGCACCGCGGCCAGGACCAACGACGA 1295							
DB	124 TGTGTGCCGAGAGGTGAAGGAGGAGCAGCACCGCGGCCAGGAGGCGACACCGACATGATGA 65							
QY	1296 TTGAGTTCGAGAGGAG 1314							

Db 64 GTCATGCTCCGACCAAGGAG 46

RESULT 13
LOCUS BP201868
DEFINITION BP201868 Sugano cDNA library, amygdala Homo sapiens EST 14-SEP-2004
ACCESSION AMR06683, mRNA sequence.
VERSION BP201868
KEYWORDS BP201868.1 GI:52052270
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 458)
Suzuki Y., Yamashita R., Shirota M., Sakakibara Y., Chiba J., Mizushima-Sugano J., Nakai K. and Sugano S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yeuzuki@ims.u-tokyo.ac.jp

FEATURES
source
1..458
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="AMR06683"
/tissue_type="amygdala"
/clone_lib="Sugano cDNA library, amygdala"

ORIGIN
Query Match 22.3%; Score 336; DB 5; Length 458;
Best Local Similarity 97.9%; Pred. No. 2.4e-88;
Matches 372; Conservative 0; Mismatches 5; Indels 3; Gaps 3;
QY 2 CAGGTTCCTGGGATGACTCTTCTCAATTGAACACTCAATTCATGGAGACCAAGAGC 61
DB 71 CAGGTTCCTGGGATGACTCTTCTCAATTGAACACTCAATTCATGGAGACCAAGAG 130
QY 62 ACTAATGCTTTGTGCTCATTCATATTTGAATCGAGGCATTGGGAACCTCTATGCTCTGT 121
DB 131 ACTAATGCTTTGTGCTCATTCATATTTGAATCGAGGCATTGGGAACCTCTATGCTCTGT 190
QY 122 TTGTGGAAGAACCAAGTGACACCACTACTGAGCTTCTCTAAAGTTCTGAAGAGTTAGAGG 181
DB 191 TTGTGGAAGAACCAAGTGACACCACTACTGAGCTTCTCTAAAGTTCTGAAGAGTTAGAGG 250
QY 182 ACTATACATCTTTCTTTTGAACCTTTTATAATAATTTATGCTCTGCG-TTTTGGGAACCCAGG 240
DB 251 ACTATACATCTTTCTTTTGAACCTTTTATAATAATAATTTATGCTCTGCGTTTGGGAACCCAGG 310
QY 241 ACTGTTAG-GGGTGAGTGACAGTCTTAC-AGTGGCCTTAATCCAACTCCAGAAATTGC 298
DB 311 GCTGTTAGAGGGGTGAGTGACAGTCTTACAGTGGGCTTATTCCAACTCCAGAAATTGC 370
QY 299 CCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCCAACTCAATCC 358
DB 371 CCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCCAACTCAATCC 430
QY 359 CTCCTTAATGATCATGATGG 378
DB 431 CTCCTTAATGATCATGATGG 450

BF963063 352 bp mRNA linear EST 22-JAN-2001
LOCUS

DEFINITION PM3-NN1174-191200-014-fl1 NN1174 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF963063
VERSION BF963063.1 GI:12380338
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 352)
Dias Neto E., Garcia Correa R., Verjovski-Almeida S., Briones M.R., Nagai M.A., da Silva W. Jr., Zago M.A., Bordini S., Costa F., Goldmann G.H., Carvalho A.F., Matsukuma A., Baia G.S., Simpson D.H., Brunstein A., de Oliveira P.S., Bucher P., Jongeneel C.V., O'Hare M.J., Soares F., Brentani R.R., Reis L.F., de Souza S.J. and Simpson A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM3&t2=PM3-NN1174-191200-014-fl1&t3=2000-12-19&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 115.
Location/Qualifiers
1..352
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN1174"
/notes="Organ: nervous normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 20.0%; Score 302; DB 4; Length 352;
Best Local Similarity 97.1%; Pred. No. 3e-78;
Matches 338; Conservative 1; Mismatches 6; Indels 3; Gaps 3;
QY 603 AGCAGTCTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATAAAAGTCAAGTGGCAAC 662
DB 4 AACAGCGGTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATAAGTCAAGTGGCAAC 63
QY 663 AGAACCTCCAGGAAAAGAAATTCGAAGAAATGAATTTAGCTGTGAGGTATGTGGCA 722
DB 64 AGAACCTCCAGGAAAAGAAATTCGAAGAAATGAATTTAGCTGTGAGGTATGTGGGC- 122
QY 723 GACATTTAGAGTCGCTTTTGTGATGTTCCAGATCCACATGAGAACACACAAAGATTCCTTCC 782
DB 123 GACATTTAGAGTCGCTTTTGTGATGTTCCAGATCCACATGAGAACACACAAAGATTCCTTCC 182
QY 783 TTACGGGTGTAAACATGTGCGGAAGAAAGATTCAGGAGCCCTTGGTTT-CTTAAAAATCACA 841
DB 183 TTACGGGTGTAAACATGTGCGGAAGAAAGATTCAGGAGCCCTTGGTTTCTTTAAAAATCACA 242
QY 842 TGGCGGACRCATTAATGCAAAATCGGGGCCGAAGCAAACTGCAAGCAAGCGCTTGAGAGTA 901
DB 243 TGGCGGACATATATGCAAAATCGGGGCCGAAGCAAACTGCAAGCAAGCGCTTGAGAGTA 302

RESULT 14
BF963063
LOCUS

QY	902	GTCCAGC-AACGATCAACGAGTCTCTCCAGGTGCACGGCGCGAGC	948
Db	303	GTCCAGCAACGATCAACGAGTCTCTCCAGGTGCACGGCGCGAGC	350

RESULT 15
 BP237807
 LOCUS BF237807 768 bp mRNA linear EST 14-NOV-2000
 DEFINITION 601842326F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4079935 5',
 mRNA sequence.

ACCESSION BF237807
 VERSION BF237807.1 GI:11151713
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 JOURNAL NIH-MGC <http://mgs.nci.nih.gov/>
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM939 row: c column: 08
 High quality sequence stop: 464.

FEATURES
 source
 1..768
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4079935"
 /tissue_type="leiomyosarcoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_46"
 /note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

ORIGIN
 Query Match 19.2%; Score 289.6; DB 2; Length 768;
 Best Local Similarity 96.7%; Pred. No. 2e-74;
 Matches 325; Conservative 2; Mismatches 6; Indels 3; Gaps 3;

QY	1175	GCATCCCTCAGTCGATCCGTTCCACACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAA	1234
Db	1	GCATCCCTCAGTCGATCCGTTCCACACCTTCCAGGCTTGGCAGCTGGCTACCAAGGAA	60
QY	1235	WAGTTGCCAT-TTGCAGAGAGTGAAGGAATTGGGCAAGAGGAGCACCACCAACGAC	1293
Db	61	AAGTTGCCATATTGCCAAGAGTGAAGGAATCGGGGCAAGAGGAGCACCACCAACGAC	120
QY	1294	GATTGAGTTCCGAGAAGGAGCTTGGAGAAACAATAAGAACCATTTGTGCAGGCTCTCG	1353
Db	121	GATTGAGTTCCGAGAAGGAGCTTGGAGAAACAATAAGGCGAGTTGTGCAGGCTCTCG	180
QY	1354	CA-AGAGAAAGAGAGTGCACACTCCACGGCGAAGCGCCCTCCGTGGAGCGGGATCC	1412
Db	181	CATAGAGAAAGAGAGTGCACACTCCACGGCGAAGCGCCCTCCGTGGAGCGGGATCC	240

QY	1413	CAAG-TTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCAGTCCGCGAAAGCTTTCA	1471
Db	241	CAGATTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCAGTCCGCGAAAGCTTTCA	300
QY	1472	GAACCTACCAACGAGCTGGTCTTTGCACTCCAGGGTCC	1507
Db	301	GAACCTACCAACGAGCTGGTCTTTGCACTCCAGGGTCC	336

Search completed: June 29, 2005, 05:14:37
 Job time : 3252.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:57:12 ; Search time 7484.35 Seconds
(without alignments)
16865.305 Million cell updates/sec

Title: US-08-731-499-4
Perfect score: 2605
Sequence: 1 CAAAGTCGAAATTAACCTC.....CTGGCGTTACCAACTTAAT 2605

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ste.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2605	100.0	2605	6	AR070327	AR070327 Sequence
2	2605	100.0	2605	6	BD085728	BD085728 Genes fro
3	1158.4	44.5	106736	9	AL354993	AL354993 Human DNA
4	1092	41.9	1122	9	BC006282	BC006282 Homo sapi
5	1076.2	41.3	1243	6	AR339439	AR339439 Sequence
6	1073	41.2	1087	6	E50594	E50594 Human prote
7	1073	41.2	1087	6	BD095367	BD095367 Human pro
8	786.8	30.2	1112	10	BC029181	BC029181 Mus muscu
9	697.6	26.8	116397	9	AL157902	AL157902 Human DNA
10	644	24.7	134465	9	AC100793	AC100793 Homo sapi
11	644	24.7	187723	9	AC016889	AC016889 Homo sapi
12	531	20.4	531	6	E50584	E50584 Human prote
13	531	20.4	531	9	AB014763	AB014763 Homo sapi
14	476	18.3	495	6	AR412671	AR412671 Sequence
15	476	18.3	495	6	AX969505	AX969505 Sequence
16	476	18.3	495	6	BD108224	BD108224 EST and e
17	453	17.4	487	6	AR412670	AR412670 Sequence
18	453	17.4	487	6	AX969504	AX969504 Sequence
19	453	17.4	487	6	BD108223	BD108223 EST and e

20	444	17.0	494	6	AR414265	AR414265 Sequence
21	444	17.0	494	6	AX971099	AX971099 Sequence
22	444	17.0	494	6	BD109818	BD109818 EST and e
23	408.4	15.7	197101	2	AC133242	AC133242 Rattus no
24	389.6	15.0	417	6	AR424010	AR424010 Sequence
25	389.6	15.0	417	6	AX984704	AX984704 Sequence
26	389.6	15.0	417	6	BD119563	BD119563 EST and e
27	386.4	14.8	400	6	AR424014	AR424014 Sequence
28	386.4	14.8	400	6	AX984708	AX984708 Sequence
29	386.4	14.8	400	6	BD119567	BD119567 EST and e
30	384.6	14.8	195444	2	AC025424	AC025424 Mus muscu
31	384.6	14.8	250353	10	AL590969	AL590969 Mouse DNA
32	362	13.9	401	6	AR424009	AR424009 Sequence
33	362	13.9	401	6	AX984703	AX984703 Sequence
34	362	13.9	401	6	BD119562	BD119562 EST and e
35	343.2	13.2	588	6	CQ715946	CQ715946 Sequence
36	328.2	12.6	383	6	AR424011	AR424011 Sequence
37	328.2	12.6	383	6	AX984705	AX984705 Sequence
38	328.2	12.6	383	6	BD119564	BD119564 EST and e
39	319.6	12.3	894	5	BC079766	BC079766 Xenopus l
40	315	12.1	983	5	BX950269	BX950269 Gallus ga
41	312.6	12.0	817	5	CR762222	CR762222 Xenopus t
42	312.6	12.0	970	5	BX930330	BX930330 Gallus ga
43	312.6	12.0	971	5	BX929492	BX929492 Gallus ga
44	312.6	12.0	999	5	BX930730	BX930730 Gallus ga
45	312.6	12.0	1001	5	CR406053	CR406053 Gallus ga

ALIGNMENTS

RESULT 1
AR070327
LOCUS AR070327 2605 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 4 from patent US 5892010.
ACCESSION AR070327
VERSION AR070327.1 GI:7221215
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2605)
AUTHORS Gray,J., Collins,C., Hwang,S.-i., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20Q13 amplicon and their uses
JOURNAL Patent: US 5892010-A 4 06-APR-1999;
FEATURES
Location/Qualifiers
source
1..2605
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN	Query Match	Best Local Similarity	Score	DB	Length
	100.0%;	100.0%;	2605;	DB 6;	2605;
	Matches 2605;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CAAGTCGAAATTAACCTCCTCACTAAAGGGAACAAAGCTGGAGCTCCACCGGCGGCGG	60		
Db	1	CAAGTCGAAATTAACCTCCTCACTAAAGGGAACAAAGCTGGAGCTCCACCGGCGGCGG	60		
QY	61	CGCTCTAGAACTAGTGATCCCGGGCTGAGGAATTCGGACGAGCTGGGCTACTAC	120		
Db	61	CGCTCTAGAACTAGTGATCCCGGGCTGAGGAATTCGGACGAGCTGGGCTACTAC	120		
QY	121	GATGCGATGAGTTCGAGTGGCGGCTGGCAGTATCGCTTCCACCTCTCTTACGTTACA	180		
Db	121	GATGCGATGAGTTCGAGTGGCGGCTGGCAGTATCGCTTCCACCTCTCTTACGTTACA	180		
QY	181	ACCGAATGTGACACTCGGCAAGCAGCTGGCGCGCTGGTGTCTGGTGTCTGTCTT	240		
Db	181	ACCGAATGTGACACTCGGCAAGCAGCTGGCGCGCTGGTGTCTGGTGTCTGTCTT	240		
QY	241	CTGCCCTGCAACAAAGTCCAGATGCGGTGATGGAAGCTCAGGAGAGCCCGCTCTT	300		

Db 241 CTGCCGCTGCACAAACAGTCCAGCATGACGCTGATGGAAGCTCAGGAGAGCCCGCTCTT 300
QY 301 CAAACAGTCAAGCTACAGCGAAAGCTTCTGTGAGTGCATCCAGATTGTTATAGAGGA 360
Db 301 CAAACAGCTCAAGCTACAGCGAAAGCTTCTGTGAGTGCATCCAGATTGTTATAGAGGA 360
QY 361 ACTGAGGAAGAAGGGAACCTCAGTGTGTTGATAGAGCAAGTCCAGCTTCTTCATCAT 420
Db 361 ACTGAGGAAGAAGGGAACCTCAGTGTGTTGATAGAGCAAGTCCAGCTTCTTCATCAT 420
QY 421 GTGGCGGAGGCCAGAGAATGGGGGAAACTCATCTATCAGTGGGTTCAGGAGTGGCCA 480
Db 421 GTGGCGGAGGCCAGAGAATGGGGGAAACTCATCTATCAGTGGGTTCAGGAGTGGCCA 480
QY 481 GAAACAACTCCGTCTTTACCTGTATGAACCTGACTGAATGGGGAAGACACAGAGGATGAGGA 540
Db 481 GAAACAACTCCGTCTTTTACCTGTATGAACCTGAATGGGGAAGACACAGAGGATGAGGA 540
QY 541 GTTCCACGGCTGGATGAAGCCACTCTACTGGGGCTCTGCAGGCCCTACAGCAGGAGCA 600
Db 541 GTTCCACGGCTGGATGAAGCCACTCTACTGGGGCTCTGCAGGCCCTACAGCAGGAGCA 600
QY 601 CAAGGCCGAGATCATCACTGTCCAGCGATGGCCGAGGCGTCAAGTTCCTTAGCAGGGACC 660
Db 601 CAAGGCCGAGATCATCACTGTCCAGCGATGGCCGAGGCGTCAAGTTCCTTAGCAGGGACC 660
QY 661 TGTCTCCCTTTACTTCTTACTCCCACTTCCAGGCTTCCAGGAGGAGACAGACCCAG 720
Db 661 TGTCTCCCTTTACTTCTTACTCCCACTTCCAGGCTTCCAGGAGGAGACAGACCCAG 720
QY 721 TGTCCCCCAAGACTGGATCTGTGACTCCACAGACTCAAAAGGAGTCCAGTCCCTGAAGG 780
Db 721 TGTCCCCCAAGACTGGATCTGTGACTCCACAGACTCAAAAGGAGTCCAGTCCCTGAAGG 780
QY 781 CTGGGACCTGGGGATGGGTTCCTACACCCCATATGTCTGTCCCTTGGATAGGGTGAAGC 840
Db 781 CTGGGACCTGGGGATGGGTTCCTACACCCCATATGTCTGTCCCTTGGATAGGGTGAAGC 840
QY 841 TGAAGCACAGGGAGAAATATGTGCTTCTCGCCCTCACTCCTTTCCCATCTCCTAGAC 900
Db 841 TGAAGCACAGGGAGAAATATGTGCTTCTCGCCCTCACTCCTTTCCCATCTCCTAGAC 900
QY 901 TGTCTCTGAGCCAGGCTGTGAAACCTGACACTTTATATGTCTCACACATGTAAGTACA 960
Db 901 TGTCTCTGAGCCAGGCTGTGAAACCTGACACTTTATATGTCTCACACATGTAAGTACA 960
QY 961 TACACACATGCGCCTGCAGCACATGCTTCTGTCTCCTCCTCCACCCCTTTAGCTGC 1020
Db 961 TACACACATGCGCCTGCAGCACATGCTTCTGTCTCCTCCTCCTCCACCCCTTTAGCTGC 1020
QY 1021 TGTTCGCTCCCTTCTCAGGCTGGTCTGTGATCCTTCTAGGGGATGGGGAGCCCTGCG 1080
Db 1021 TGTTCGCTCCCTTCTCAGGCTGGTCTGTGATCCTTCTAGGGGATGGGGAGCCCTGCG 1080
QY 1081 TGCAGGACGCTTCCAGGCAATATGAAGATAGGAGCCACAGGGCTGCGAGTGAAGGT 1140
Db 1081 TGCAGGACGCTTCCAGGCAATATGAAGATAGGAGCCACAGGGCTGCGAGTGAAGGT 1140
QY 1141 GTGGCCCCACACCGATTATGATATATAAATCTCAACTCCCAAAAAAAGGGGAGAGGT 1200
Db 1141 GTGGCCCCACACCGATTATGATATATAAATCTCAACTCCCAAAAAAAGGGGAGAGGT 1200
QY 1201 CTGAGACTAGTCTCTCTCTCGAGAACTAGTCTCGAGTTCCTTTTCTTTTCTTTT 1260
Db 1201 CTGAGACTAGTCTCTCTCTCGAGAACTAGTCTCGAGTTCCTTTTCTTTTCTTTT 1260
QY 1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1320
Db 1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1320
QY 1321 ACTTTTCTCTACTTAATACTACTTTCCAGTCTCAGAAGCCAGAGGGGAAAAAAGAC 1380

Db 1321 ACTTTTCTCTACTTAATACTACTTTTCCAGTCTCAGAAGCCAGAGGGGAAAAAAGAC 1380
QY 1381 CATGAATCTTCTCTCCAGATTAAAGTACACACTTTTGGAAAAACAGATTGGAAACCTTT 1440
Db 1381 CATGAATCTTCTCTCCAGATTAAAGTACACACTTTTGGAAAAACAGATTGGAAACCTTT 1440
QY 1441 CTGAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTTGATGTTGCTCATGAA 1500
Db 1441 CTGAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTTGATGTTGCTCATGAA 1500
QY 1501 AATTGTTAAAAACCTGTTCTAGATAAAGAACAGTCTCAAGTTTTTGTACAGCCTACACAT 1560
Db 1501 AATTGTTAAAAACCTGTTCTAGATAAAGAACAGTCTCAAGTTTTTGTACAGCCTACACAT 1560
QY 1561 AGTACAAGGGTCCCTATGATGATTCCTCTGTAGGACGAATAATGTAATTTTTTCAGTT 1620
Db 1561 AGTACAAGGGTCCCTATGATGATTCCTCTGTAGGACGAATAATGTAATTTTTTCAGTT 1620
QY 1621 TCTGGTTTTATAACTCTCTCGATCTCAGAGTTGACTGATTTAAAAACACCTACTCATGCAACA 1680
Db 1621 TCTGGTTTTATAACTCTCTCGATCTCAGAGTTGACTGATTTAAAAACACCTACTCATGCAACA 1680
QY 1681 GAGATAAAGCACTCATATTTTTTATAAATATATAGGACCAAACTATTTTTGGAAATCTTAT 1740
Db 1681 GAGATAAAGCACTCATATTTTTTATAAATATATAGGACCAAACTATTTTTGGAAATCTTAT 1740
QY 1741 CTATTGGAGACACATATGCTGGACTAAAGCAATAATTTATTTCTCAATGTCGTGCG 1800
Db 1741 CTATTGGAGACACATATGCTGGACTAAAGCAATAATTTATTTCTCAATGTCGTGCG 1800
QY 1801 TAACTCTCAATGACTTAGAATGCTTTGCTATATTTTTGCTCTATGCTCTCAACACACTGCG 1860
Db 1801 TAACTCTCAATGACTTAGAATGCTTTGCTATATTTTTGCTCTATGCTCTCAACACACTGCG 1860
QY 1861 TTTCTTTTAGCTCTTTGAAACAAGCCTGCTCTGCTCAGGACCAAGATTTTGGGA 1920
Db 1861 TTTCTTTTAGCTCTTTGAAACAAGCCTGCTCTGCTCAGGACCAAGATTTTGGGA 1920
QY 1921 CTTCTCTTAAGATTCTATTTCTTAATTTCTTAATCTGGGTAACTTAGTTTATCCAACA 1980
Db 1921 CTTCTCTTAAGATTCTATTTCTTAATTTCTTAATCTGGGTAACTTAGTTTATCCAACA 1980
QY 1981 CTTAGATCTCTGCCGTAATAAACTCTTTCTATAGAGCCCTGTCATGACACTGCTCTCTTC 2040
Db 1981 CTTAGATCTCTGCCGTAATAAACTCTTTCTATAGAGCCCTGTCATGACACTGCTCTCTTC 2040
QY 2041 TCCAAACATACCTCAACAGACACATGTAAGTACAGTCTGCTCTGTTTTCTTTTCA 2100
Db 2041 TCCAAACATACCTCAACAGACACATGTAAGTACAGTCTGCTCTGTTTTCTTTTCA 2100
QY 2101 TACTTTTCTCTATCATGCTTCCCTCCATTATAATTTTTTATATGTTGTAATGTCGTG 2160
Db 2101 TACTTTTCTCTATCATGCTTCCCTCCATTATAATTTTTTATGTTGTTGTAATGTCGTG 2160
QY 2161 CCCCAAGTCAGTTTCTCTCACTAAACTATAAACTCCGTAAGCTGGGATCTTCCCAATTT 2220
Db 2161 CCCCAAGTCAGTTTCTCTCACTAAACTATAAACTCCGTAAGCTGGGATCTTCCCAATTT 2220
QY 2221 GATCACCACTTAGTACAGTAGGAAACACAGTAAAGATTCAATTTGGTATTTTGTGGAATGAT 2280
Db 2221 GATCACCACTTAGTACAGTAGGAAACACAGTAAAGATTCAATTTGGTATTTTGTGGAATGAT 2280
QY 2281 GAAATGAATTTTGTCTAGTAAAGTCTGGGGGAAACCCAGGTGAGAGAGCCCTAGAAAGCA 2340
Db 2281 GAAATGAATTTTGTCTAGTAAAGTCTGGGGGAAACCCAGGTGAGAGAGCCCTAGAAAGCA 2340
QY 2341 GGTTCGAATCCAAGGCTAGTAGACTTAGTGTACTCAAGAAGGGGTAGCCTGAAATAAA 2400
Db 2341 GGTTCGAATCCAAGGCTAGTAGACTTAGTGTACTCAAGAAGGGGTAGCCTGAAATAAA 2400
QY 2401 GGTTCGAATTTATGTCGAAGATAGTCAAGACATGCGGCAAGACAGAGTGTCTGTCGCC 2460
Db 2401 GGTTCGAATTTATGTCGAAGATAGTCAAGACATGCGGCAAGACAGAGTGTCTGTCGCC 2460

QY 2461 GAATTCGATATCAAGCTTATTCGATATCGTCCGACCTCGAGGGGGGGCCCCGGTACCCAAATTC 2520
DB 2461 GAATTCGATATCAAGCTTATTCGATATCGTCCGACCTCGAGGGGGGGCCCCGGTACCCAAATTC 2520
QY 2521 GCCCTATAGTGGTCTGATTTACAAATTCACCTGCGCGTCTGTTTACACAGTCTGACTGGGA 2580
DB 2521 GCCCTATAGTGGTCTGATTTACAAATTCACCTGCGCGTCTGTTTACAAACGCTGACTGGGA 2580
QY 2581 AAACCTGGCGTTTACCCTTAAT 2605
DB 2581 AAACCTGGCGTTTACCCTTAAT 2605

RESULT 2
BD085728
LOCUS BD085728 2605 bp DNA linear PAT 27-AUG-2002
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085728
VERSION BD085728.1 GI:22631338
KEYWORDS JP 2001524802-A/4.
SOURCE synthetic construct
ORGANISM other sequences: artificial sequences.
REFERENCE 1 (bases 1 to 2605)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 4 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
COMMENT OS Artificial Sequence
PN JP 2001524802-A/4
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998056264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI DAVID KOWBEL,
PI JOHANNA ROMMENS
CC C12N15/11,C12Q1/68,A61K48/00
PC Description of Artificial Sequence:cc43
FH Key Location/Qualifiers
FT source 1..2605
FT Location/Qualifiers
1..2605 /organism='Artificial Sequence'.
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'

FEATURES
source
Query Match 100.0%; Score 2605; DB 6; Length 2605;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCTCGAAATTAACCTTCACCTAAAGGGAACAAAAGCTGGAGCTCCACCGGTGGCGG 60
DB 1 CAAGCTCGAAATTAACCTTCACCTAAAGGGAACAAAAGCTGGAGCTCCACCGGTGGCGG 60
QY 61 CGCTCTAGAACTAGTGGATCCCGGGCTGCGAGGAATTCGGACAGCTGGGCTACTAC 120
DB 61 CGCTCTAGAACTAGTGGATCCCGGGCTGCGAGGAATTCGGACAGCTGGGCTACTAC 120
QY 121 GATGGCGATGAGTTTCAGTGGCGGCGGAGTATCCCTCCACCCCTCTTTACGTTTACA 180
DB 121 GATGGCGATGAGTTTCAGTGGCGGCGGAGTATCCCTCCACCCCTCTTTACGTTTACA 180
QY 181 ACCGAAATGGACACCTCGGCAAGACAGCTGGCGGCTGGTCTGCTGGTCTGTCCTT 240
DB 181 ACCGAAATGGACACCTCGGCAAGACAGCTGGCGGCTGGTCTGCTGGTCTGTCCTT 240
QY 241 CTGCGCCCTGCAAAACAGTCCAGCATGACGGTATGGAAGCTCAGGAGAGCCCGCTCTT 300

DB 241 CTGCGCCCTGCAAAACAGTCCAGCATGACGGTATGGAAGCTCAGGAGAGCCCGCTCTT 300
QY 301 CAACAAACGTCAAGCTACAGCGAAAGCTTCTCTGTGGAGTGCATCCAGATTTGTTATGAGGA 360
DB 301 CAACAAACGTCAAGCTACAGCGAAAGCTTCTCTGTGGAGTGCATCCAGATTTGTTATGAGGA 360
QY 361 ACTGAGGAAGAAAGGAACCTCGAGTGGTGGATGATAGAGCAAGTCCAGCTTCTGTATCAT 420
DB 361 ACTGAGGAAGAAAGGAACCTCGAGTGGTGGATGATAGAGCAAGTCCAGCTTCTGTATCAT 420
QY 421 GTGGCGGAGGCCAGAAAGTGGGGAACCTCATCTATCATGTGGGTTTCCAGGAGTGGCCA 480
DB 421 GTGGCGGAGGCCAGAAAGTGGGGAACCTCATCTATCATGTGGGTTTCCAGGAGTGGCCA 480
QY 481 GAACAACTCCGTCTTTACCTGTATGAACCTGATAATGGGGAAGACACAGAGGATGAGGA 540
DB 481 GAACAACTCCGTCTTTACCTGTATGAACCTGATAATGGGGAAGACACAGAGGATGAGGA 540
QY 541 GTTCCACGGGCTGGATGAAGCCACTCTACTCGGGCTCTGCAGGCCCTTACAGCAGGAGCA 600
DB 541 GTTCCACGGGCTGGATGAAGCCACTCTACTCGGGCTCTGCAGGCCCTTACAGCAGGAGCA 600
QY 601 CAAGGCCGAGATCATCACTGTCTAGCGATGGCGAGCGCTCAAGTTCTTCTAGCAGGAGCC 660
DB 601 CAAGGCCGAGATCATCACTGTCTAGCGATGGCGAGCGCTCAAGTTCTTCTAGCAGGAGCC 660
QY 661 TGTCTCCCTTTACTTCTTTACCTCCACCTTTCCAGGGCTTTCAAAGGAGACAGACCCAG 720
DB 661 TGTCTCCCTTTACTTCTTTACCTCCACCTTTCCAGGGCTTTCAAAGGAGACAGACCCAG 720
QY 721 TGTCTCCCTTTACTTCTGTGATCTCACAGACTCAAAAGGACTCAGTCTCTGAAAGG 780
DB 721 TGTCTCCCTTTACTTCTGTGATCTCACAGACTCAAAAGGACTCAGTCTCTGAAAGG 780
QY 781 CTGGGACCTGGGAGTGGGTTTCTCACACCCCATATGTCTGTCTCTTGGATAGGCTCAGGC 840
DB 781 CTGGGACCTGGGAGTGGGTTTCTCACACCCCATATGTCTGTCTCTTGGATAGGCTCAGGC 840
QY 841 TGAAGCACCGAGGAGAAAATATGTCTCTCTCGCCCTACCTCTCTTCCCATCTCTAGAC 900
DB 841 TGAAGCACCGAGGAGAAAATATGTCTCTCTCTCGCCCTACCTCTCTTCCCATCTCTAGAC 900
QY 901 TGTCTCTGAGCCAGGGTCTGTAAACCTTGACACTTTTATATGTGTTCACATGTAAGTACA 960
DB 901 TGTCTCTGAGCCAGGGTCTGTAAACCTTGACACTTTTATATGTGTTCACATGTAAGTACA 960
QY 961 TACACACATGCGCCTGCGACACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
DB 961 TACACACATGCGCCTGCGACACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
QY 1021 TGTTCGCTCCCTTCTCAGGCTGGTGGATCTTCTCTAGGGGATGGGGGAAGCCCTGGC 1080
DB 1021 TGTTCGCTCCCTTCTCAGGCTGGTGGATCTTCTCTAGGGGATGGGGGAAGCCCTGGC 1080
QY 1081 TCGAGGAGCCCTTTCAGGCAATATGAAGATAGAGGCCCAACCGGCTCTGCGAGTGAAGGT 1140
DB 1081 TCGAGGAGCCCTTTCAGGCAATATGAAGATAGAGGCCCAACCGGCTCTGCGAGTGAAGGT 1140
QY 1141 GTGGCCCCACCGGATTTATGATATTTAAATCTCACTCCCAAAAAAAGGAAAAA 1200
DB 1141 GTGGCCCCACCGGATTTATGATATTTAAATCTCACTCCCAAAAAAAGGAAAAA 1200
QY 1201 CTGAGACTAGTCTCTCTCTCTCGAGAACTAGTCTCGAGTCTTCTTCTTCTTCTTCTTCT 1260
DB 1201 CTGAGACTAGTCTCTCTCTCTCGAGAACTAGTCTCGAGTCTTCTTCTTCTTCTTCTTCT 1260
QY 1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 1320
DB 1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 1320
QY 1321 ACTTTTCTTCTACTTAACTACTTTTCTCGAGTCTCAGAGCCCGAGGAAAAAAGAAC 1380
DB 1321 ACTTTTCTTCTACTTAACTACTTTTCTCGAGTCTCAGAGCCCGAGGAAAAAAGAAC 1380

QY 1381 CATGAATCTTCTCCAGATTAAAGTACACACTTTTGGAAAAAGATTTGGAAAAACCTTTT 1440
 Db 1381 CATGAATCTTCTCTCCAGATTAAAGTACACACTTTTGGAAAAAGATTTGGAAAAACCTTTT 1440
 QY 1441 CTGAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTTGATGTTGCTCATGAA 1500
 Db 1441 CTGAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTTGATGTTGCTCATGAA 1500
 QY 1501 AATTGTTAAAAACCTGTTCTAGATAAAGAACAGTCTCAAGTTTGTACAGCCTACACAT 1560
 Db 1501 AATTGTTAAAAACCTGTTCTAGATAAAGAACAGTCTCAAGTTTGTACAGCCTACACAT 1560
 QY 1561 AGTACAAGGTCCTCTATGATGATCTTCTGTAGACGAAATAATGTAATTTTTTCAGTT 1620
 Db 1561 AGTACAAGGTCCTCTATGATGATCTTCTGTAGACGAAATAATGTAATTTTTTCAGTT 1620
 QY 1621 TCTGTTTTATAACTCTCTCGATCTCAGAGTTGACTGATTTAAACACCTACTCATGCAACA 1680
 Db 1621 TCTGTTTTATAACTCTCTCGATCTCAGAGTTGACTGATTTAAACACCTACTCATGCAACA 1680
 QY 1681 GAGATAAAGCACTCATATTTTATAAATTATATGACCAAACTATTTTGGAAATCTTAT 1740
 Db 1681 GAGATAAAGCACTCATATTTTATAAATTATATGACCAAACTATTTTGGAAATCTTAT 1740
 QY 1741 CTATTGGAGACAAATATGCTGCACTAAAGCAATAATTTATTTCTCAATGCTGTGTC 1800
 Db 1741 CTATTGGAGACAAATATGCTGCACTAAAGCAATAATTTATTTCTCAATGCTGTGTC 1800
 QY 1801 TAACTCAATGACTTTAGATGCTTTGCTATATTTTGCCTCTATGCTCAACACACTGCGC 1860
 Db 1801 TAACTCAATGACTTTAGATGCTTTGCTATATTTTGCCTCTATGCTCAACACACTGCGC 1860
 QY 1861 TTTCTTTTGTCTTGAACAGCCAACTGCTTCTGCTCAGGACGAGATATTTTGGGA 1920
 Db 1861 TTTCTTTTGTCTTGAACAGCCAACTGCTTCTGCTCAGGACGAGATATTTTGGGA 1920
 QY 1921 CTTCTCTTAAAGAAATCTATTTCTTAAATCTTTATCTGGTAACTAGTTTATCCAACA 1980
 Db 1921 CTTCTCTTAAAGAAATCTATTTCTTAAATCTTTATCTGGTAACTAGTTTATCCAACA 1980
 QY 1981 CTTGAGATCTGCGCGTAAAACTCTTCTTATAGAGCCGTGTCATGACACTGCTCTCTTC 2040
 Db 1981 CTTGAGATCTGCGCGTAAAACTCTTCTTATAGAGCCGTGTCATGACACTGCTCTCTTC 2040
 QY 2041 TCCACATACTCACAGACACACATGACTAGACTAGATAGAACCTCTCTGTTTCTTTTCA 2100
 Db 2041 TCCACATACTCACAGACACACATGACTAGACTAGATAGAACCTCTCTGTTTCTTTTCA 2100
 QY 2101 TACTTTTCTCTATCATGCTTCCCTCCATATATAATTTTATATGTTGTAATGTCGT 2160
 Db 2101 TACTTTTCTCTATCATGCTTCCCTCCATATATAATTTTATATGTTGTAATGTCGT 2160
 QY 2161 CCCCAAGTCAGTTTCTCTCACTAACTATAAACTCCGTAAGCTGGATCCTTCCAAATTT 2220
 Db 2161 CCCCAAGTCAGTTTCTCTCACTAACTATAAACTCCGTAAGCTGGATCCTTCCAAATTT 2220
 QY 2221 GATCACCTTAGTACAGTGAACACAGTAAGATTCATTTGTTGTAATGTAAT 2280
 Db 2221 GATCACCTTAGTACAGTGAACACAGTAAGATTCATTTGTTGTAATGTAAT 2280
 QY 2281 GAATCAATTTGTTTCTAGTAAAGTCTGGGGAAACCCAGGTGAGAAGAGCCTAGAAAGCA 2340
 Db 2281 GAATCAATTTGTTTCTAGTAAAGTCTGGGGAAACCCAGGTGAGAAGAGCCTAGAAAGCA 2340
 QY 2341 GGTGGAATCCAGGCTAGATAGACTTAGTGTGTTACTCAAGAAAGGCTAGCTGAAATAAA 2400
 Db 2341 GGTGGAATCCAGGCTAGATAGACTTAGTGTGTTACTCAAGAAAGGCTAGCTGAAATAAA 2400
 QY 2401 GGTTCAAATATTAGTCAAGAAATAGTCAAGACATGGGCAAGACAGTGTCTGCTGCTGCC 2460
 Db 2401 GGTTCAAATATTAGTCAAGAAATAGTCAAGACATGGGCAAGACAGTGTCTGCTGCTGCC 2460

QY 2461 GAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCGGTACCAATTC 2520
 Db 2461 GAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCGGTACCAATTC 2520
 QY 2521 GCCCTATAGTGAGTCGTATTACAAATTCATCTGGCGCTGCTTTTACAACTGTCGTGACTGGGA 2580
 Db 2521 GCCCTATAGTGAGTCGTATTACAAATTCATCTGGCGCTGCTTTTACAACTGTCGTGACTGGGA 2580
 QY 2581 AAACCCCTGGCGTTACCCCAACTTAAT 2605
 Db 2581 AAACCCCTGGCGTTACCCCAACTTAAT 2605

RESULT 3
 AL354993

LOCUS AL354993 106736 bp DNA linear PRI 20-JUN-2001

DEFINITION Human DNA sequence from clone RP5-823G15 on chromosome 20q13.2-13.3. Contains a peptidylprolyl isomerase A (cyclophilin A) (PPIA) pseudogene, the gene for OVC10-2, part of a gene encoding a putative zinc finger protein ZNF218, ESTs, STSs and GSSs, complete sequence.

ACCESSION AL354993

VERSION AL354993.24 GI:1121037

KEYWORDS HTG; cyclophilin; OVC10-2; peptidoprolyl isomerase; PPIA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 106736)

AUTHORS Frankland, J.

TITLE Direct Submission

JOURNAL Submitted (15-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

COMMENT requests: clonerequest@sanger.ac.uk

On Nov 8, 2000 this sequence version replaced gi:11096473.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

IMPORTANT: This sequence is not the entire insert of clone RP5-823G15. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP4-724E16 is at 106637 in this sequence. The true right end of clone RP4-669H2 is at 100 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-823G15 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

FEATURES

Location/Qualifiers

1..106736

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="20"

/map="q13.2-13.3"

```

/clone="RP5-823G15"
/clone_lib="RPCI-5"
complement(1607..2252)
/notes="match: GSS: Em:AQ547110"
/notes="match: GSS: Em:AQ547110"
complement(2016..2175)
/notes="match: GSS: Em:AQ544248"
complement(2281..2716)
/notes="match: GSS: Em:AQ214808"
complement(4163..4653)
/notes="match: GSS: Em:AQ214808"
complement(4295..4728)
/notes="match: GSS: Em:AQ39057"
complement(4329..4727)
/notes="match: GSS: Em:AQ452763"
4728..5164
/notes="match: GSS: Em:AQ505606"
complement(4944..5321)
/notes="match: GSS: Em:AQ586033"
10396..10768
/notes="Sequence from clone PCR only."
10697..10880
/notes="match: GSS: Em:AQ345020"
complement(10704..10882)
/notes="match: GSS: Em:AQ480874 Em:AQ484778"
complement(10725..10890)
/notes="match: GSS: Em:AQ410161"
10769
/notes="Tandem repeat. Forced join. Gap size estimated to
be approximately 250bp by restriction digest data."
10865..10878
/notes="Single clone region. Assembly confirmed by
restriction digest data. Sequence from uni-directional
dntp read only."
10872..10901
/notes="15 copies 2 mer ac 96% conserved"
12294..12331
/notes="29 copies 2 mer ac 94% conserved"
join(12456..12685,16697..17055)
/notes="match: STS: Em:G31647"
16264..16395
/notes="match: GSS: Em:AQ18306"
16286..16731
/genes="dJ823G15.1"
/pseudo
16286..16731
/genes="dJ823G15.1"
/notes="dJ823G15.1 (peptidylprolyl isomerase A (cyclophilin
A) (PPIA) pseudogene)
match: proteins: Tr:Q62523 Tr:Q9Y536 Sw:P17742 Sw:P30405
Tr:Q9U0H1 Sw:P14851 Tr:Q9ZNU7 Sw:P29117 Sw:P04374
Sw:P05092 Sw:P10111 Tr:Q9TTC6 Sw:P25007 Tr:Q65104
Sw:P54985 Sw:P52011 Tr:P52015 Sw:P21568 Tr:Q40672
Sw:P23284 Sw:P32011 Sw:P34790 Sw:Q09637 Tr:Q41119
Sw:P24367 Sw:P43369 Tr:P73789 Sw:O00060 Sw:P26882
Sw:P14832 Sw:P30414 Sw:P30415 Sw:P18253 Tr:O88541
Tr:O93970 Sw:P45877 Sw:P80311 Tr:O74729 Tr:O22515
Sw:Q41651"
/pseudo
/codon_start=1
/evidence=not_experimental
16399..16570
/genes="dJ823G15.1"
/notes="match: GSS: Em:AQ18306"
/pseudo
complement(join(16562..16727,17042..17121))
/notes="match: GSS: Em:AQ043681"
19639..20229
/notes="L1PA10 repeat: matches 5563..6165 of consensus"
complement(21939..22096)
/notes="match: GSS: Em:AL303319"
complement(22725..23122)
/notes="match: GSS: Em:AQ005428"
27061..27754

```

```

repeat_region
27757..27811
/notes="L1PB3 repeat: matches 5446..6150 of consensus"
28635..28694
/notes="30 copies 2 mer ac 95% conserved"
36251..36326
/notes="38 copies 2 mer tt 69% conserved"
37995..38120
/notes="match: STS: Em:G43859"
42888..42984
/notes="LOR1a repeat: matches 399..497 of consensus"
43241..43654
/notes="LOR1a repeat: matches 2..399 of consensus"
47872..47975
/notes="match: STS: Em:HSPF16G4"
49776..50267
/notes="LOR1a repeat: matches 1..496 of consensus"
50646..55870
/notes="L1PA2 repeat: matches 900..6144 of consensus"
55866..56638
/notes="L1PA2 repeat: matches 1..776 of consensus"
complement(57366..57862)
/notes="match: GSS: Em:B32842"
60846..61136
/notes="match: GSS: Em:B41668"
60846..61038
/notes="match: GSS: Em:AQ197282"
63180..63585
/notes="match: GSS: Em:AQ293265"
63216..63649
/notes="match: GSS: Em:AQ568074"
63759..63790
/notes="16 copies 2 mer aa 90% conserved"
64118..64163
/notes="23 copies 2 mer ta 95% conserved"
64615..64889
/notes="match: GSS: Em:AQ749930"
64674..64923
/notes="125 copies 2 mer ga 62% conserved"
complement(64679..65039)
/notes="match: GSS: Em:AQ236314"
complement(65694..66150)
/notes="match: GSS: Em:AQ590175"
complement(65898..66178)
/notes="match: GSS: Em:AQ028660"
67717..68496
/notes="L1PA2 repeat: matches 1..776 of consensus"
68492..73743
/notes="L1PA2 repeat: matches 900..6146 of consensus"
complement(76682..76998)
/notes="match: STS: Em:HSPF17C6"
complement(78885..79361)
/notes="match: GSS: Em:AQ680262"
complement(80657..80886)
/notes="match: GSS: Em:B38354"
complement(80671..81110)
/notes="match: GSS: Em:AQ698368"
82216..82551
/genes="dJ823G15.3"
82216..82551
/genes="dJ823G15.3"
/product="dJ823G15.3 (Novel protein possibly zinc finger
protein 218 ZNF218 (reported in Literature (PNAS vol
95:8703-8708)))"
/notes="Continues in Em:AC004505
match: ESTs: Em:AA406371"
/evidence=not_experimental
complement(83637..83894)
/notes="match: STS: Em:G24374"

```

Query Match 44.5%; Score 1158.4; DB 9; Length 106736;
 Best Local Similarity 99.7%; Pred. No. 1.2e-253;
 Matches 1171; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1279 TGGCTTTAAGGATTTATTTATTTGTTTCCCTTTTACAGTGTCCACTTTTCTCTACTTAATA 1338
Db 45960 TGGCTTTAAGGATTTATTTA--GTTTCCCTTTTACAGTGTCCACTTTTCTCTACTTAATA 46017
QY 1339 CTACTTTTCAGTCTCAGAACCCAGAGGGGAAAAAAGACACATGAAATCTTCTCTCC 1398
Db 46018 CTACTTTTCAGTCTCAGAACCCAGAGGGGAAAAAAGACACATGAAATCTTCTCTCC 46077
QY 1399 AGATTAAAGTACACACTTTTGGAACACAGATTGGAAACCTTTTCTGAAAAAGTTGACTGA 1458
Db 46078 AGATTAAAGTACACACTTTTGGAACACAGATTGGAAACCTTTTCTGAAAAAGTTGACTGA 46137
QY 1459 TACTCCAAACCAACATGCAATGCTGATGTTGCTCATGAAATTTGTTAAAAAAGTTT 1518
Db 46138 TACTCCAAACCAACATGCAATGCTGATGTTGCTCATGAAATTTGTTAAAAAAGTTT 46197
QY 1519 CTAGATAAAGAACAGTCTCAAGTTTTTGTACAGCCTTACACATAGTACAGGGTCCCTAT 1578
Db 46198 CTAGATAAAGAACAGTCTCAAGTTTTTGTACAGCCTTACACATAGTACAGGGTCCCTAT 46257
QY 1579 GATGATCTCTGTAGGACGAATAATGTAATTTTTCAGTTTCTGGTTTATAACTCTCT 1638
Db 46258 GATGATCTCTGTAGGACGAATAATGTAATTTTTCAGTTTCTGGTTTATAACTCTCT 46317
QY 1639 CGATCTCAGATTGACTGATTAACACCTACTCATGCAACAGAGAAATAAGACCTCAT 1698
Db 46318 CGATCTCAGATTGACTGATTAACACCTACTCATGCAACAGAGAAATAAGACCTCAT 46377
QY 1699 TTTTATATAATATATGACCAAACTATTTTGGAAATCTTATCTATGAGACACAAAT 1758
Db 46378 TTTTATATAATATATGACCAAACTATTTTGGAAATCTTATCTATGAGACACAAAT 46437
QY 1759 GCTGACTAAAGCAATAATATTTATCTCAATGCTGTGTCTAATCACTCAATGACTTGA 1818
Db 46438 GCTGACTAAAGCAATAATATTTATCTCAATGCTGTGTCTAATCACTCAATGACTTGA 46497
QY 1819 ATGCTTTGCTATATTTTGCCTATGCTCAACACACAGCTGCTTTCTTTAGCTCTTGA 1878
Db 46498 ATGCTTTGCTATATTTTGCCTATGCTCAACACACAGCTGCTTTCTTTAGCTCTTGA 46557
QY 1879 CAAGCAAACTGCTTCCCTCAGGACACAGATATTTTGGGACTTCTCTTGAAGATTTCT 1938
Db 46558 CAAGCAAACTGCTTCCCTCAGGACACAGATATTTTGGGACTTCTCTTGAAGATTTCT 46617
QY 1939 TTTCCCTTAATCTTATCTGGGTAACTTAGTTTTATCCAACTTCAGATCCGCGTAA 1998
Db 46618 TTTCCCTTAATCTTATCTGGGTAACTTAGTTTTATCCAACTTCAGATCCGCGTAA 46677
QY 1999 AAACCTCTTATAGAGCCTGTCATGACACTGCTCTCTTCCAAACATATCTACCCAGC 2058
Db 46678 AAACCTCTTATAGAGCCTGTCATGACACTGCTCTCTTCCAAACATATCTACCCAGC 46737
QY 2059 ACACATGTAGACTAGATTAGAACCTTCTGTTTTCTTTTTCATCTTTTCTATCATGC 2118
Db 46738 ACACATGTAGACTAGATTAGAACCTTCTGTTTTCTTTTTCATCTTTTCTATCATGC 46797
QY 2119 TTCCCTCCATTAATAATTTTATATGTTGTTGAATGTTCTGCCCAAGTCAAGTTTCTCTC 2178
Db 46798 TTCCCTCCATTAATAATTTTATATGTTGTTGAATGTTCTGCCCAAGTCAAGTTTCTCTC 46857
QY 2179 ACTTAACTATAAATCCGTTAAAGCTGGGATCTTCCAAATTTTGTACACCACTTAGTACAG 2238
Db 46858 ACTTAACTATAAATCCGTTAAAGCTGGGATCTTCCAAATTTTGTACACCACTTAGTACAG 46917
QY 2239 TAGGAACACAGTAAAGATTCAATTTGGTATTTTGTGGAATGAATGAATTTGTTTGCTA 2298
Db 46918 TAGGAACACAGTAAAGATTCAATTTGGTATTTTGTGGAATGAATGAATTTGTTTGCTA 46977
QY 2299 GTAAAGTCTGGGGGAAACCCAGTGTAGAGAGCCTTGAAGAGAGGTGCAATCCAGGCTAG 2358
Db 46978 GTAAAGTCTGGGGGAAACCCAGTGTAGAGAGCCTTGAAGAGAGGTGCAATCCAGGCTAG 47037

QY 2359 ATAGACTTAGTGTCTACTCAAGAAAGGCTAGCTGAAATAAAAGTTTCAAAATATATAGTCAA 2418
Db 47038 ATAGACTTAGTGTCTACTCAAGAAAGGCTAGCTGAAATAAAAGTTTCAAAATATATAGTCAA 47097
QY 2419 GAATAGTCAAGACATGGGCAAGACAGAGAGTGCTG 2452
Db 47098 GAATAGTCAAGACATGGGCAAGACAGAGAGTGCTG 47131

RESULT 4
BC006282
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BC006282
Homo sapiens hypothetical protein MGC10540, mRNA linear PRI 30-JUN-2004
MGC:10540 IMAGE:4025884, complete cds.
BC006282
BC006282.2 GI:34785072
MGC.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1122)
Straussberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.P., Jordan,H., Moore,T., Wax,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schneitz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Sutcliffe,A., Schein,J.E., Jones,S.J., Skalska,U., Smalau,D.E.,
Schnerf,A., Schein,J.E., Jones,S.J., and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1122)
Straussberg,R.
Direct Submission
Submitted (09-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Sep 16, 2003 this sequence version replaced gi:136223362.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nih.gov
Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Madsen,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Shantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 15 Row: d Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 14150154.

FEATURES
 Location/Qualifiers
 source
 1..1122
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:10540 IMAGE:4025884"
 /tissue type="Skin, melanotic melanoma."
 /clone lib="NIH MGC 20"
 /lab_host="DH10B-R"
 /note="Vector: pOT87"
 1..1122
 /gene="MGC10540"
 /db_xref="LocusID:84313"
 41..571
 /gene="MGC10540"
 /codon_start=1
 /product="hypothetical protein MGC10540"
 /protein_id="AAH06282.1"
 /db_xref="GI:13623363"
 /db_xref="LocusID:84313"
 /translation="MAMSEFWQVRPPFFLQPNVDTRQKLAACWCLVLSFRLH
 KQSMVMAQSPFNKKLPVSIQVLEELKKGNLEWLDKSKSFLIMWR
 RPEWKLTYQVRSRGSQNNVFTLYELTNGBDTEDEPHGLDEATLLRALQALQEH
 KABIITVSDGRGVKRP"
 ORIGIN

Query Match 41.9%; Score 1092; DB 9; Length 1122;
 Best Local Similarity 100.0%; Pred. No. 8.5e-239;
 Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 CTGGGCTACTAGCGATGATGTTTCAGTGGCGCGTGGCAGTATCGCTTCCACCCCTT 168
 DB 28 CTGGGCTACTAGCGATGATGTTTCAGTGGCGCGTGGCAGTATCGCTTCCACCCCTT 87

QY 169 CTTTACGTTACAAACCGAATGTGACACTCGGCAGAGAGAGCTGGCGCTGTGCTCGCT 228
 DB 88 CTTTACGTTACAAACCGAATGTGACACTCGGCAGAGAGAGCTGGCGCTGTGCTCGCT 147

QY 229 GGTCTGTCTTCTGCGGCTGCACAAACAGTCCAGCATGAGTGTGATGGAAGCTCAGGA 288
 DB 148 GGTCTGTCTTCTGCGGCTGCACAAACAGTCCAGCATGAGTGTGATGGAAGCTCAGGA 207

QY 289 GAGCCGCTCTTCAACAGCTCAAGCTACAGCGAAGCTTCTGTGGAGTCCGATCCAGAT 348
 DB 208 GAGCCGCTCTTCAACAGCTCAAGCTACAGCGAAGCTTCTGTGGAGTCCGATCCAGAT 267

QY 349 TGTATTAGAGGAACCTGAGGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAG 408
 DB 268 TGTATTAGAGGAACCTGAGGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAG 327

QY 409 CTTCTGTATCATGTGGCGAGGCCAGAGAATGGGGGAACTCATCTATCAGTGGGTTTC 468
 DB 328 CTTCTGTATCATGTGGCGAGGCCAGAGAATGGGGGAACTCATCTATCAGTGGGTTTC 387

QY 469 CAGGAGTGGCCAGAACTCCGCTTTTACCTGTATGACTGACTTAATGGGGAGACAC 528
 DB 388 CAGGAGTGGCCAGAACTCCGCTTTTACCTGTATGACTGACTTAATGGGGAGACAC 447

QY 529 AGAGGATGAGGAGTTCACAGGGCTGATGAAGCCACTTACTACGGGGCTCTGACGGCCCT 588
 DB 448 AGAGGATGAGGAGTTCACAGGGCTGATGAAGCCACTTACTACGGGGCTCTGACGGCCCT 507

QY 589 ACAGCAGGAGCAACAGCCAGATCATCTATGTCAGCATGAGCCAGGCGTCAAGTTCTT 648
 DB 508 ACAGCAGGAGCAACAGCCAGATCATCTATGTCAGCATGAGCCAGGCGTCAAGTTCTT 567

QY 649 CTAGCAGGAGCACTGTCTTCTTACTTCTTACTCCCACTTTTCCAGGGCTTCAAGAG 708
 DB 568 CTAGCAGGAGCACTGTCTTCTTACTTCTTACTCCCACTTTTCCAGGGCTTCAAGAG 627

QY 709 AGACAGACCCAGTGTCTCCCAAGAGACTGTGACTTCCACAGACTCAAAAGGACTC 768
 DB 628 AGACAGACCCAGTGTCTCCCAAGAGACTGTGACTTCCACAGACTCAAAAGGACTC 687

QY 769 CAGTCTGAGGCTGGGAGCTGGGATGGTTCACACCCCATATGTCTGTCCCTTGG 828
 DB 688 CAGTCTGAGGCTGGGAGCTGGGATGGTTCACACCCCATATGTCTGTCCCTTGG 747

QY 829 ATAGGCTGAGGCTGAAGACACAGGGAGAGAAATATGTGCTTCTTCGCGCTTACTCCTTT 888
 DB 748 ATAGGCTGAGGCTGAAGACACAGGGAGAGAAATATGTGCTTCTTCGCGCTTACTCCTTT 807

QY 889 CCATCTTAGACTGTCTTGTAGCCAGGGTCTGTAAACCTGACACTTATATGTGTTTACA 948
 DB 808 CCATCTTAGACTGTCTTGTAGCCAGGGTCTGTAAACCTGACACTTATATGTGTTTACA 867

QY 949 CATGTAAGTACATACACATGCGCTGCAGACATGCTTCTGTCTCTCTCTCTCTCTCCAC 1008
 DB 868 CATGTAAGTACATACACATGCGCTGCAGACATGCTTCTGTCTCTCTCTCTCTCTCCAC 927

QY 1009 CCTTTAGCTGTGTCT 1068
 DB 928 CCTTTAGCTGTGTCT 987

QY 1069 GGAAGCCCTGTGTGAGGAGCTTCCAGGCAATATGAAGATAGGAGGCCACGGGCTG 1128
 DB 988 GGAAGCCCTGTGTGAGGAGCTTCCAGGCAATATGAAGATAGGAGGCCACGGGCTG 1047

QY 1129 GCAGTGAGAGGTGTGGCCCCACACCGATTTATGATATTAATAATCTCAACTCCCAAAAA 1188
 DB 1048 GCAGTGAGAGGTGTGGCCCCACACCGATTTATGATATTAATAATCTCAACTCCCAAAAA 1107

QY 1189 AAAAAAATAA 1200
 DB 1108 AAAAAAATAA 1119

RESULT 5
 LOCUS AR339439 1243 bp DNA linear PAT 17-AUG-2003
 DEFINITION Sequence 930 from patent US 6569662.
 ACCESSION AR339439
 VERSION AR339439.1 GI:33726296
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 UNCLASSIFIED.
 REFERENCE 1 (bases 1 to 1243)
 AUTHORS Tang, Y.T., Zhou, P. and Drmanac, R.T.
 TITLE Nucleic acids and polypeptides
 JOURNAL Patent: US 6569662-A 930 27-MAY-2003;
 FEATURES
 Location/Qualifiers
 source
 1..1243
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 41.3%; Score 1076.2; DB 6; Length 1243;
 Best Local Similarity 99.7%; Pred. No. 3.5e-235;
 Matches 1078; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 113 GCTCTACGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCCCTTCTTT 172
 DB 163 GCGTCGACGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCCCTTCTTT 222

QY 173 AGCTTACAAACCGAATGTGACACTCGGCAGAGAGAGCTGGCCGCTGTGCTGTGCTGTGTC 232
 DB 223 AGCTTACAAACCGAATGTGACACTCGGCAGAGAGAGCTGGCCGCTGTGCTGTGCTGTGTC 282

QY 233 CTGTCTTCTGCGGCTTGCACAAACAGTCCAGCATGCGGTATGGAAGCTCAGGAGAGC 292
 DB 283 CTGTCTTCTGCGGCTTGCACAAACAGTCCAGCATGCGGTATGGAAGCTCAGGAGAGC 342

QY	293	CGCTCTTCAACAACTCAAGCTACAGCGAAAGCTTCTGTGAGTCGATCCAGATTGTA	352
Db	343	CGCTCTTCAACAACTCAAGCTACAGCGAAAGCTTCTGTGAGTCGATCCAGATTGTA	402
QY	353	TTAGAGGAACCTGAGGAAGAAAGGAACCTCGATGGTTGGATAAGAGCAAGTCCAGCTTC	412
Db	403	TTAGAGGAACCTGAGGAAGAAAGGAACCTCGATGGTTGGATAAGAGCAAGTCCAGCTTC	462
QY	413	CTGATCATGTGCGGAGGCGCAGAGAATGGGGAAACTCATCTATCATGTTGGTTCCAGG	472
Db	463	CTGATCATGTGCGGAGGCGCAGAGAATGGGGAAACTCATCTATCATGTTGGTTCCAGG	522
QY	473	AGTGCCCAAGAACAACTCCGTCTTTTACCTGTATGAACCTGACTAAATGGGGAAGACACAG	532
Db	523	AGTGCCCAAGAACAACTCCGTCTTTTACCTGTATGAACCTGACTAAATGGGGAAGACACAG	582
QY	533	GATGAGGAGTTCACGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGCGAGGCCCTACAG	592
Db	583	GATGAGGAGTTCACGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGCGAGGCCCTACAG	642
QY	593	CAGGAGCAAGGCGGAGATCATCTGTACCGATGGCGGAGGCGTCAAGTCTTCTAG	652
Db	643	CAGGAGCAAGGCGGAGATCATCTGTACCGATGGCGGAGGCGTCAAGTCTTCTAG	702
QY	653	CAGGACCTGTCTCCCTTTTACTTCTTACTCTCCCTCCACTTTTCCAGGGCTTTCAAAGGAGAC	712
Db	703	CAGGACCTGTCTCCCTTTTACTTCTTACTCTCCCTCCACTTTTCCAGGGCTTTCAAAGGAGAC	762
QY	713	AGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGT	772
Db	763	AGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGT	822
QY	773	CCTGAAGGCTGGGACTCGGGATGGGTTTCTCACACCCCATATGTCTGTCCTTGGATAG	832
Db	823	CCTGAAGGCTGGGACTCGGGATGGGTTTCTCACACCCCATATGTCTGTCCTTGGATAG	882
QY	833	GCTGAGGCTGAAGCACACAGGGAGAAATATGTGCTTCTCGCCCTTACCTCCTTTCCCA	892
Db	883	GCTGAGGCTGAAGCACACAGGGAGAAATATGTGCTTCTCGCCCTTACCTCCTTTCCCA	942
QY	893	TCCTAGACTGTCTTTGAGCCAGGGTCTGTAAACCTGACACTTTATATGTTGTTTCAACATG	952
Db	943	TCCTAGACTGTCTTTGAGCCAGGGTCTGTAAACCTGACACTTTATATGTTGTTTCAACATG	1002
QY	953	TAAGTACATACACATATGGGCTGACGACATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT	1012
Db	1003	TAAGTACATACACATATGGGCTGACGACATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT	1062
QY	1013	TTAGTGTCTGTGCT	1072
Db	1063	TTAGTGTCTGTGCT	1122
QY	1073	GCCCTGGCTGACGAGGCTTCCAGGCAATATGAAGATAGGAGGCCCAACGGGCTTGGCAG	1132
Db	1123	GCCCTGGCTGACGAGGCTTCCAGGCAATATGAAGATAGGAGGCCCAACGGGCTTGGCAG	1182
QY	1133	TGAGAGGTGTGCCCCCACACCGATTATGATTTAAATCTCAACTCCCAAAAAA	1192
Db	1183	TGAGAGGTGTGCCCCCACACCGATTATGATTTAAATCTCAACTCCCAAAAAA	1242
QY	1193	A	1193
Db	1243	A	1243
RESULT 6			
E50594			
LOCUS	E50594	1087 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Human protein and cDNA [3].		
ACCESSION	E50594		
VERSION	E50594.1 GI:22555017		
KEYWORDS	JP 2001161368-A/14.		

SOURCE	Homo sapiens (human)									
ORGANISM	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE	1 (bases 1 to 1087)									
AUTHORS	Kato,S. and Saeki,M.									
TITLE	Human protein and cDNA [3]									
JOURNAL	Patent: JP 2001161368-A 14 19-JUN-2001;									
	SCIENCE & TECH AGENCY									
COMMENT	OS Homo sapiens (human)									
	PN JP 2001161368-A/14									
	PD 19-JUN-2001									
	PF 06-DEC-1999 JP 1999346864									
	PI SEIJI KATO,MIHORO SAEKI									
	PC									
	C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/									
	10//									
	PC C12P21/02,C12P21/08,C12N15/00,C12N5/00									
CC										
FH	Key	Location/Qualifiers								
FT	CDS	(28)..(558).								
FEATURES	source									
	1..1087									
	/organism="Homo sapiens"									
	/mol_type="genomic DNA"									
	/db_xref="taxon:9606"									
ORIGIN										
	Query Match	41.2%; Score 1073; DB 6; Length 1087;								
	Best Local Similarity	100.0%; Pred. No. 1.9e-234;								
	Matches 1073; Conservative	0; Mismatches 0; Indels 0; Gaps 0;								
QY	109	CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGCGGTGGCAGTATCGCTTCCACCCCTT	168							
Db	15	CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGCGGTGGCAGTATCGCTTCCACCCCTT	74							
QY	169	CTTTACGTTTAAACCGAATGTGGACACTCGGCAGAAAGCAGCTGGCGCCCTGGTGTCTCGCT	228							
Db	75	CTTTACGTTTAAACCGAATGTGGACACTCGGCAGAAAGCAGCTGGCGCCCTGGTGTCTCGCT	134							
QY	229	GGTCTGTCTTCTGCGGCTGCACAAAACAGTCCAGCATGACGGTGTGGAAGCTCAGGA	288							
Db	135	GGTCTGTCTTCTGCGGCTGCACAAAACAGTCCAGCATGACGGTGTGGAAGCTCAGGA	194							
QY	289	GAGCCGCTCTTCAACACGCTCAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGAT	348							
Db	195	GAGCCGCTCTTCAACACGCTCAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGAT	254							
QY	349	TGTATTAGAGGAACGTAGGAAGAAAGGAACTTCGAGTGGTTGGATAAGAGCAAGTCCAG	408							
Db	255	TGTATTAGAGGAACGTAGGAAGAAAGGAACTTCGAGTGGTTGGATAAGAGCAAGTCCAG	314							
QY	409	CTTCTGTGATCATGTGGCGGAGGCGCAGAGAATGGGGGAAACTCATCTATCAGTGGGTTTC	468							
Db	315	CTTCTGTGATCATGTGGCGGAGGCCAGAGAATGGGGGAAACTCATCTATCAGTGGGTTTC	374							
QY	469	CAGAGTGGCCAGAACAACTCCGTCTTTTACCTCTGTATGAACCTGACTAATGGGGAAGACAC	528							
Db	375	CAGAGTGGCCAGAACAACTCCGTCTTTTACCTCTGTATGAACCTGACTAATGGGGAAGACAC	434							
QY	529	AGAGGATGAGGAGTTTCCACGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGAGAGCCCT	588							
Db	435	AGAGGATGAGGAGTTTCCACGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGAGAGCCCT	494							
QY	589	ACAGCAGGAGCACAAAGGCCGAGATCATCATCTGTACGCGATGGCCGAGGCGTCAAGTCTTT	648							
Db	495	ACAGCAGGAGCACAAAGGCCGAGATCATCATCTGTACGCGATGGCCGAGGCGTCAAGTCTTT	554							
QY	649	CTAGCAGGACCTGTCTCTCTTTTACTTCTTACTCTCCACCTTTCAGGGGCTTTCAAAGG	708							
Db	555	CTAGCAGGACCTGTCTCTCTTTTACTTCTTACTCTCCACCTTTCAGGGGCTTTCAAAGG	614							
QY	709	AGACAGACCCAGTGTCTCCCAAGAAAGACTGGATCTGTGACTTCCACAGACTCAAAAGGACTC	768							

Db 1035 GCAGTGAGAGGTGTGGCCCCCACACCGATTATGATATTAATAATCTCAACTCCC 1087

RESULT 8
BC029181
LOCUS
DEFINITION
MUS musculus DNA segment, Chr 11, Wayne State University 68,
expressed, mRNA (cdna clone MGC:35627 IMAGE:4240500), complete cds.
ACCESSION
BC029181
VERSION
BC029181.1 GI:20809732
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1112)
Strausberg,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettner,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalek,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1112)
Strausberg,R.
Direct Submission
Submitted (01-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleeged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 61 Row: m Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 25092661.
Location/Qualifiers
1. .1112
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:35627 IMAGE:4240500"
/tissue_type="Kidney, normal, 5 month old male mouse."

FEATURES
source

gene

/clone.lib="NCI CGAP_Kid14"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .1112
/gene="D11Wsu68e"
/db_xref="LocusID:28084"
/db_xref="MGI:106354"
13. -543
/gene="D11Wsu68e"
/codon_start=1
/product="DNA segment, Chr 11, Wayne State University 68,
expressed"
/protein_id="AAH29181.1"
/db_xref="GI:20809733"
/db_xref="LocusID:28084"
/db_xref="MGI:106354"
/translation="MAMSFEPWQYRPFPTLPQNVDTROKQLAAMCSLVLSFRLH
KQSMVTMEAESPLFNVKLQKLPVESIOIVLEELKKGNLEWLKDKKSFLLMWR
RPEWGLKIYQWVSRSQNNVSFTLYELTSGEDTEDEFHGLDEATLLRALQALQOEH
KAEIITVSDGRGVKFF"

CDS

ORIGIN

Query Match 30.2%; Score 786.8; DB 10; Length 1112;
Best Local Similarity 87.9%; Pred. No. 5.7e-169;
Matches 963; Conservative 0; Mismatches 117; Indels 16; Gaps 9;

Qy 110 TGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGTGCAGTATCGCTTCCACCCCTTC 169
Db 1 TGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGTGCAGTATCGCTTCCGCGCTTC 60

Qy 170 TTACGTTTAAACCGAATGTGGACACTCGCGAGAGCAGCTGGCGCGCTGGTGTCTCGCTG 229
Db 61 TTACGTTTAAACCGAATGTGGACACTCGCGAGAGCAGCTGGCGCGCTGGTGTCTCTG 120

Qy 230 GTCTGTCTTCTGCGGCTTGCACAAACAGTCCAGTATGAGTATGAGTATGAGTATGAGT 289
Db 121 GTCTGTCTTCTGCGGCTTGCACAAACAGTCCAGTATGAGTATGAGTATGAGTATGAGT 180

Qy 290 AGCGCGCTTCAACACGTCAGCTACAGGAGAACTTCCTGTGAGTTCGATCGATTCAGATT 349
Db 181 AGCGCGCTTCAACACGTCAGCTACAGGAGAACTTCCTGTGAGTTCGATTCAGATT 240

Qy 350 GTATTAGAGGAACCTGAGGAAGAAAGGAAACCTCGAGTGGTGGTATGAGTATGAGTATG 409
Db 241 GTATTAGAGGAACCTGAGGAAGAAAGGAAACCTCGAGTGGTGGTATGAGTATGAGTATG 300

Qy 410 TTCCTGATCATGTGGCGAGGCGCAGAGAAATGGGGGAAACTCATCTATCAGTGGGTTTCC 469
Db 301 TTCCTGATCATGTGGCGAGGCGCAGAGAAATGGGGGAAACTCATCTATCAGTGGGTTTCC 360

Qy 470 AGGAGTGGCCAGACAACTCCGCTTTTACCTGTATGAACCTGACTTAATGGGGAAGACACA 529
Db 361 AGGAGTGGCCAGATTAATCTCTGTGTTTACTCTGTATGAACCTGACTTAATGGGGAAGACACA 420

Qy 530 GAGGATGAGGAGTTCCACGGGCTGGATGAAGCCTCTATCTGCGGGCTCTGAGGCCCTTA 589
Db 421 GAGGATGAAGAGTTCCATGGGCTGGACGAGGCGACCTCTGCGGCGCTCTGAGGCCCTTA 480

Qy 590 CAGCAGGAGCACAAGGCGGAGATCATCTGTACGATGGCGGAGGCGTCAAGTTCCTTC 649
Db 481 CAGCAGGAGCACAAGGCGTGAATATCATCTGTACGATGGCGGAGGCGTCAAGTTCCTTC 540

Qy 650 TAGCAGGACCTGTCTCCCTTTACTCTTCTTACTCTCCACCTTTCAGGGCTTTCAAAAGGA 709
Db 541 TAGCAGACCTTGGCT-CCTTTACTCTTCTTACTCTCCACCTTTCAGGGCTTTCAAAAGGA 599

Qy 710 GACGAGCCAGTGTCCCAAGAGCTGGATCTGTGATCTCCACCACTCCACCACTCAAAAGGACTCC 769
Db 600 GACGAGCCTAGTGT-CCCCCAGACTGGATCTGTCTCACTCTCTCAGACTC-AAAGGGCTCA 657

Qy 770 AGTCTGAAAGCTGGGAGCTGGGATGGGTTTCTCA-CACCCCATATGTCTGTCTCTGG 828
Db 658 AGTCTGAAAGCTCAGACTAGGGATGTTTCCACATAACCCCTTTATTCTGTCTCTGG 717

/evidence=not experimental
/product="d3675C20.1 (mannosidase, alpha, class 1A, member 2)"
/protein_id="CAC10451.1"
/db_xref="GI:10443041"
/db_xref="GOA:O60476"
/translation="HTSVGLGDSFYVLLKAWLMSDKTDHEARKWYDDAIEATEKHL
IKSRGGTLFTGEWKNHLEKMGHLACFAGGMPALGADGRADKAGHYLGLGARIAR
TKESYDRTALKLGPESPKFGQAEVAVRQAEKYILRPVIEYWLWLPRTDPRY
RQWHEAALAEIKYCRVNGSGVDYSSTPHDDVQSQFLAETLKYLLFLSGDD
LLPLDHWVFNEAHLPLVHLANTLTLGNPAVR"
7043. .7347
repeat_region
/note="AluSx repeat: matches 3. .307 of consensus"
7506. .7733
repeat_region
/note="MIR repeat: matches 1. .262 of consensus"
8136. .8179
repeat_region
/note="L2 repeat: matches 2657. .2700 of consensus"
8363. .8480
repeat_region
/note="L2 repeat: matches 2584. .2710 of consensus"
complement(10218. .10720)
misc_feature
/note="match: GSS: Em:AQ664370"
11953. .12065
repeat_region
/note="L2 repeat: matches 2592. .2695 of consensus"
complement(13408. .13965)
misc_feature
/note="match: GSS: Em:AQ310773"
13564. .14014
misc_feature
/gene="MAN1A2"
/note="match: GSS: Em:AQ876874"
15669. .15813
repeat_region
/note="MIR repeat: matches 93. .259 of consensus"
16746. .16903
repeat_region
/note="MERSA repeat: matches 13. .175 of consensus"
17228. .17336
repeat_region
/note="L2 repeat: matches 2627. .2749 of consensus"
17869. .18611
repeat_region
/note="L1M4 repeat: matches 7182. .7960 of consensus"
19003. .19046
repeat_region
/note="11 copies 4 mer at 81% conserved"
19168. .19221
repeat_region
/note="2 copies 27 mer 100% conserved"
20219. .20318
repeat_region
/note="L2 repeat: matches 2418. .2521 of consensus"
20642. .20926
repeat_region
/note="HAL1 repeat: matches 467. .763 of consensus"
24238. .24350
repeat_region
/note="L2 repeat: matches 2629. .2746 of consensus"
24497. .24864
repeat_region
/note="THE1C repeat: matches 1. .371 of consensus"
24894. .25015
repeat_region
/note="FLAM_C repeat: matches 2. .123 of consensus"
25042. .25148
repeat_region
/note="L2 repeat: matches 2621. .2742 of consensus"
25214. .25516
repeat_region
/note="L2 repeat: matches 2440. .2748 of consensus"
25877. .26061
repeat_region
/note="MERSA repeat: matches 1. .188 of consensus"
26291. .26338
repeat_region
/note="24 copies 2 mer ac 87% conserved"
26293. .26340
repeat_region
/note="12 copies 4 mer acac 87% conserved"
26549. .26738
repeat_region
/note="L1M2 repeat: matches 5903. .6120 of consensus"
26799. .26887
repeat_region
/note="L1M3 repeat: matches 6087. .6175 of consensus"
26892. .27031
repeat_region
/note="L1M2 repeat: matches 5749. .5896 of consensus"
27623. .27930
repeat_region
/note="AluSx repeat: matches 1. .309 of consensus"
28659. .28943
repeat_region
/note="MER2 repeat: matches 1. .345 of consensus"
29210. .29489
repeat_region
/note="L1M2 repeat: matches 5841. .6166 of consensus"

misc_feature complement(29909. .30339)
/note="match: GSS: Em:AQ379387"
30087. .31494
repeat_region
/note="L1M3c repeat: matches 274. .1745 of consensus"
32097. .34934
repeat_region
/note="L1PB1 repeat: matches 2354. .5224 of consensus"
34905. .35728
repeat_region
/note="L1PB1 repeat: matches 5336. .6155 of consensus"
35756. .36102
repeat_region
/note="L1M4 repeat: matches 3703. .4100 of consensus"
36401. .36820
misc_feature
/gene="MAN1A2"
/note="match: GSS: Em:AQ093391"
Query Match 26.8%; Score 697.6; DB 9; Length 116397;
Best Local Similarity 80.3%; Pred. No. 3.3e-148;
Matches 925; Conservative 0; Mismatches 199; Indels 28; Gaps 8;
QY 96 AATTCCGACGAGCTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATC 155
DB 62931 AAAGTGGGATTTCCCTGGGCTATTATGATGGTGACGAATTTCAAGTGGTTGTGGCAGTATC 62990
QY 156 GCTTCCCAACCTTTCTTTACGTTTAAACCGAATGTGGACACTCGGAGAGAGAGCTGGCGG 215
DB 62991 GCTTTCCGCGC--CTTTACATTACAGCTGAACGTGGCCACTTTGGCAGAAGCAGCTGGCCA 63048
QY 216 CTTGTGTCTCGTGTCTGTCTCTTCGCGCTGCACAAACAGTCCAGCATGACGCTGA 275
DB 63049 CTTGTGTGTGTGTGTCTGTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 63108
QY 276 TGGAGCTCCAGGAGGCGGCTCTTCAACAACGTCGCAAGCTACAGCGAAAGCTTCCTGTGG 335
DB 63109 TGGATGCTCAGGAGATCTGCTCTTCAGCAACATCAAGCT--GTGGAAGCTTCCTGTGG 63165
QY 336 AGTCGATCCAGATTGTATTAGAGGAAGTCTAGAGAAAGGAAGCACTCGAGTGGTGGATA 395
DB 63166 GATCAATCCAGTTGTATTAGAGGAAGTCTAGAGAAAGTGGGAAGTACAGTGGCTGGATA 63225
QY 396 AGAGCAAGTCCAGCTTCCTGTATGTCGCGAGGCCAGAGAAATGGGGAAGCTCATCT 455
DB 63226 AGAGCAAGTCTAGTTTCTCTTAATCATGTGGCGAGGCCAGAGAAATGGGGAAGCTCATCT 63285
QY 456 ATCAGTGGGTTTCCAGGAGTGGCCAGAAACAATCCGCTCTTTTACCCCTGTATGAAGTACT 515
DB 63286 ATCAGTGGGTTTCCAGGAGTGGCCAGAAACAATCCGCTCTTTAGCTGTATGAGTGAACCA 63345
QY 516 ATGGGAAAGACACAGAGATGAGGATTCACGGGCTGGATGAAGCACTTACTTGGCGG 575
DB 63346 ATGGGAAAGACATAGAGAAATGAGGTGTTTCCACGGAATAAAGGAGGC---CTTCTGTGG 63401
QY 576 CTCTGAGGCGCTTACAGCAGGAGCACAAAGGCCGAGATCATCACTGT-----C 622
DB 63402 CTCTGAGGCGCTTACAGTAGGAACATAAGCTGAGATCATCACTCATCTCAGGAGACC 63461
QY 623 AGCGATGGCCGAGGCGTCAAGTTCTTTAGCAGGAGACCTGTCTCCCTTTTACTTTTACTCT 682
DB 63462 AGTGTGCTGAGGTGTCAGGTCCATCTGTCTGAGACCTGTCTCTTTTACTTTTCTATCT 63521
QY 683 CCCACCTTTCCAGGCGCTTTCAAAGGAGACAGACCCAGTGTCCCCCAAGACTGATCTG 742
DB 63522 CCCACCTTTCCAGGCGCTTTCCAGGAGGAGACAGACTTCGGGTG-GAACCACACATTTGAACCTG 63580
QY 743 TGACTCCACACAGACTCAAAAGGACTCCAGTCTGAAGCTGGGAGCTGGGATGGTTTC 802
DB 63581 TGAATCCACACAGACTCAAAAGGCTCCAGTCTGAAGCTGGGAGCCAGGGGTGGGTTTC 63640
QY 803 TCACACCCCATATGTCTGTCCCTTGGATAGGCTGAGGCTGAAGCACACGAGGAGAAAAAT 862
DB 63641 TCATACACAACATGTCTGTCTCCCTGGGACAGGCTGAGGCTGAGGCACTCCACAGAAAAATG 63700
QY 863 GTGCTTCTTCTC-GCCCTACCTCTCTTTCCATCTCTAGACTGTCTTGAGCCAGGCTGTGT 921
DB 63701 CTTCTTCAAAGCTTGGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 63760

repeat_region /rpt_family="AluJb"
2400..2547
repeat_region /rpt_family="L2"
2613..2661
repeat_region /rpt_family="(TA)n"
complement(2670..2789)
repeat_region /rpt_family="FLAM_C"
complement(2795..3122)
repeat_region /rpt_family="AluY"
6180..6232
repeat_region /rpt_family="L2"
complement(6233..6530)
repeat_region /rpt_family="AluY"
6531..6585
repeat_region /rpt_family="L2"
complement(6586..6887)
repeat_region /rpt_family="AluSc"
6888..7159
repeat_region /rpt_family="L2"
7875..7912
repeat_region /rpt_family="GC_rich"
7957..8044
repeat_region /rpt_family=" "(CGGGG)n"
9590..9691
repeat_region /rpt_family="MER91A"
10084..10105
repeat_region /rpt_family="GC_rich"
10488..10589
repeat_region /rpt_family="MIR"
12437..12588
repeat_region /rpt_family="MIR3"
12868..12900
repeat_region /rpt_family=" "(CCCCCT)n"
13333..13370
repeat_region /rpt_family="GC_rich"
14420..14475
repeat_region /rpt_family="GA_rich"
14721..14744
repeat_region /rpt_family="GC_rich"
14801..14840
repeat_region /rpt_family="GC_rich"
complement(15362..15663)
repeat_region /rpt_family="AluJo"
15881..15924
repeat_region /rpt_family="GC_rich"
16110..16162
repeat_region /rpt_family=" "(CGGGG)n"
complement(16380..16658)
repeat_region /rpt_family="L3b"
complement(16694..16909)
repeat_region /rpt_family="L3"
17073..17108
repeat_region /rpt_family="GC_rich"
17379..17409
repeat_region /rpt_family="GC_rich"
18578..18710
repeat_region /rpt_family="MIR"
complement(18775..19088)
repeat_region /rpt_family="AluSx"
19780..19886
repeat_region /rpt_family="GA_rich"
19935..20031
repeat_region /rpt_family="MIR"
20334..20445
repeat_region /rpt_family="GA_rich"
20855..20960
repeat_region /rpt_family=" "(CTGGGG)n"
20906..21085
repeat_region /rpt_family=" "(TTGGGG)n"
21037..21216
repeat_region /rpt_family=" "(CTGGGG)n"
21166..21309
repeat_region /rpt_family=" "(TTGGGG)n"

repeat_region complement(22213..22416)
repeat_region /rpt_family="L2"
22417..22720
repeat_region /rpt_family="AluSg"
23332..23457
repeat_region /rpt_family="MIR"
complement(23477..23601)
repeat_region /rpt_family="MIR"
24030..24157

Query Match 24.7%; Score 644; DB 9; Length 134465;
Best Local Similarity 100.0%; Pred.No. 6.2e-136;
Matches 644; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 GAGTTCACGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGCAGGCCCTCAGCAGGAG 598
DB 116729 GAGTTCACGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGCAGGCCCTCAGCAGGAG 116788
QY 599 CACAAGGCCGAGATCATCATCTGTCTAGCGATGCCCGAGGCGTCAAGTTCTTCTAGCAGGGA 658
DB 116789 CACAAGGCCGAGATCATCATCTGTCTAGCGATGCCCGAGGCGTCAAGTTCTTCTAGCAGGGA 116848
QY 659 CTTGTCTCCCTTTACTTCTTACCTCCACCTTTCCAGGGCTTTCAAAAGGAGACAGACC 718
DB 116849 CTTGTCTCCCTTTACTTCTTACCTCCACCTTTCCAGGGCTTTCAAAAGGAGACAGACC 116908
QY 719 AGTGTCCCCCAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCTCTGAA 778
DB 116909 AGTGTCCCCCAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCTCTGAA 116968
QY 779 GCTGGGACCTGGGATGGGTTTCTCACACCCCATATGTCTGTCTCTTGGATAGGGTGAG 838
DB 116969 GCTGGGACCTGGGATGGGTTTCTCACACCCCATATGTCTGTCTCTTGGATAGGGTGAG 117028
QY 839 GCTGAAGCACCGAGGAGAAAATATGTGCTTCTTCGCGCTTACCTCTTCCATCTCTAG 898
DB 117029 GCTGAAGCACCGAGGAGAAAATATGTGCTTCTTCGCGCTTACCTCTTCCATCTCTAG 117088
QY 899 ACTGTCTTTGAGCCAGGCTCTGTAACCTGACACTTTATATGTGTTCACACATGTAAGTA 958
DB 117089 ACTGTCTTTGAGCCAGGCTCTGTAACCTGACACTTTATATGTGTTCACACATGTAAGTA 117148
QY 959 CATACACATGCGGCTGCAGCACAATGCTTCTGTCTCTCTCTCCCTCCACCCCTTTAGCT 1018
DB 117149 CATACACATGCGGCTGCAGCACAATGCTTCTGTCTCTCTCTCCCTCCACCCCTTTAGCT 117208
QY 1019 GCTGTTGCTCTCTCTTCAGGCTGGTGTGATCTTCTTAGGGATGGGGGAAGCCCTG 1078
DB 117209 GCTGTTGCTCTCTCTTCAGGCTGGTGTGATCTTCTTAGGGATGGGGGAAGCCCTG 117268
QY 1079 GCTGCAGGCAGCTTCCAGGCAATATGAAGATAGAGGCCCAACGGGCTGGCAGTGAGAG 1138
DB 117269 GCTGCAGGCAGCTTCCAGGCAATATGAAGATAGAGGCCCAACGGGCTGGCAGTGAGAG 117328
QY 1139 GTGTGGCCCCACACCGATTTATGATTTAAATCTCAACTCCCA 1182
DB 117329 GTGTGGCCCCACACCGATTTATGATTTAAATCTCAACTCCCA 117372

RESULT 11
AC016889
LOCUS AC016889 187723 bp DNA linear PRI 01-JAN-2003
DEFINITION Homo sapiens chromosome 17, clone RP11-506G7, complete sequence.
ACCESSION AC016889
VERSION AC016889.28 GI:27452920
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 187723)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone RP11-506G7

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 187723)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavskiy, L., Boukhgalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 187723)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (24-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 187723)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (01-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 1, 2003 this sequence version replaced gi:25229251.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information

Center project name: L5105
Center clone name: 506_g_7

Only the last 187.8 kilobases of this clone are being submitted.
The remainder overlaps accession number AC100793 [WICGR project L20710].

FEATURES
source

Location/Qualifiers
1..187723
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-506G7"
/clone_lib="RPC1-11 Human Male BAC"
89..321
/rpt_family="AluJb"
322..337
/rpt_family="GA-rich"
338..648
/rpt_family="AluSx"
649..685
/rpt_family="GA-rich"
769..937
/rpt_family="AluSx"
1272..1331
/rpt_family="L1ME1"
1332..1628
/rpt_family="AluSg"
1629..2166
/rpt_family="L1ME1"
2179..2477
/rpt_family="AluSx"
complement(2615..2911)
/rpt_family="AluSx"
2945..3253
/rpt_family="AluJb"
3521..3823
/rpt_family="AluY"
4035..4092
/rpt_family="T-rich"
complement(4111..4146)
/rpt_family="Alu"
complement(4149..4235)
/rpt_family="AluSc"
complement(4236..4377)
/rpt_family="AluSg/x"
complement(4382..4681)
/rpt_family="AluSc"
5013..5033
/rpt_family="AT-rich"
5076..5363
/rpt_family="AluSg"
complement(5415..5621)
/rpt_family="AluY"
complement(5676..5869)
/rpt_family="AluJc"
complement(5870..6161)
/rpt_family="AluSc"
complement(6162..6273)
/rpt_family="AluJc"
complement(6859..7124)
/rpt_family="AluSx"
complement(7317..7614)
/rpt_family="AluJb"
complement(7768..8068)
/rpt_family="AluSg"
complement(8250..8413)
/rpt_family="L2"
complement(8498..8794)
/rpt_family="AluSx"
complement(8798..9096)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:28:52 ; Search time 911.543 Seconds
(without alignments)
16917.381 Million cell updates/sec

Title: US-08-731-499-4
Perfect score: 2605
Sequence: 1 CAACTCGAATTAACCTC.....CTGGCGTTACCACTTAAT 2605

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2604.6	100.0	2605	AAV04699	Aav04699 Homo sapi
2	1089	41.8	1089	AAA40516	Aaa40516 Murine ad
3	1076.2	41.3	1243	AAI59037	Aai59037 Human pol
4	1076.2	41.3	1243	ADQ99260	Adq99260 DNA encod
5	1076.2	41.3	1243	ADB49020	Adb49020 Novel hum
6	1075	41.3	1098	AAI60823	Aai60823 Human pol
7	1073	41.2	1087	AAH68547	Aah68547 Human pro
8	1073	41.2	1100	ADQ84631	Adq84631 Human tum
9	603.6	23.2	634	ABQ57032	Abq57032 Human col
10	515.8	19.8	559	ABQ56941	Abq56941 Human col
11	489.2	17.8	507	ABQ56838	Abq56838 Human can
12	464.4	17.8	507	ABQ56838	Abq56838 Human col
13	395	15.2	609	ABQ56915	Abq56915 Human col
14	353.6	13.6	545	ABX17514	Abx17514 Human cdn
15	317.4	12.2	349	ABQ56970	Abq56970 Human col
16	269	10.3	269	ABV89089	Abv89089 Human col
17	269	10.3	269	ABV88953	Abv88953 Human col
18	269	10.3	269	ABK39342	Abk39342 DNA encod
19	269	10.3	269	ACA11671	Aca11671 Human lung
20	269	10.3	269	ACA02857	Aca02857 Lung canc

C 21	269	10.3	269	10	ADH46899	Adh46899 Human lun
C 22	269	10.3	269	13	ADJ20818	Adj20818 Human lun
C 23	260	10.0	260	6	ABV86907	Abv86907 Human col
C 24	253.6	9.7	490	10	ADB56763	Adb56763 Toxicity-
25	253	9.7	304	3	AAA43467	Aaa43467 Mouse sec
26	238	9.1	276	2	AAV26672	Aav26672 Mouse nov
27	227	8.7	241	3	AAZ42852	Aaz42852 Human 5'
C 28	219.2	8.4	469	5	ABV37980	Abv37980 Human pro
C 29	218.2	8.4	335	5	ABV25168	Abv25168 Human pro
30	190	7.3	195	3	AAC00435	Aac00435 Human sec
C 31	187	7.2	409	5	AAF66957	Aaf66957 Novel hum
32	177.8	6.8	270	12	ADP28751	Adp28751 Human sec
C 33	169.4	6.5	217	4	AAI84882	Aai84882 Human pol
C 34	169.4	6.5	217	9	ACH31848	Ach31848 Human bon
35	166.6	6.4	246	2	AAV26673	Aav26673 Mouse nov
C 36	163.6	6.3	230	8	ABX38439	Abx38439 Bovine ES
C 37	156.4	6.0	1763	2	AAZ30709	Aaz30709 Rat neuro
C 38	156.4	6.0	1827	2	AAZ28264	Aaz28264 Rat neuro
C 39	154.2	5.9	2201	4	AAZ26054	Aaz26054 Human cdn
C 40	154.2	5.9	2201	8	ABX73395	Abx73395 Human nov
C 41	152.6	5.9	3315	5	ABA14281	Abai4281 Human ner
C 42	146.2	5.6	8710	2	AAV32370	Aav32370 Complete
C 43	146.2	5.6	8710	3	AAAS9039	Aas9039 Nucleotid
C 44	146.2	5.6	8710	6	ABA94248	Abas94248 Adenoviru
C 45	146.2	5.6	8710	10	ADB75145	Adb75145 Plaemid p

ALIGNMENTS

RESULT 1
AAV04699
ID AAV04699 standard; CDNA; 2605 BP.
XX
AC AAV04699;
XX
DT 21-JUL-1998 (first entry)
XX
DE Homo sapiens 20q13 amplicon cc43 transcript.

XX 20q13 amplicon; chromosome 20; tumour; detection;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentata;
KW Leber's congenital amaurosis; ds.

OS Homo sapiens.
XX
PN WO9802539-A1.

XX
PD 22-JAN-1998.

XX
PP 15-JUL-1997; 97WO-US012343.

XX
PR 15-JUL-1996; 96US-00680395.

XX
PR 16-OCT-1996; 96US-00731499.

XX
PR 17-JAN-1997; 97US-00785532.

XX
PA (REGC) UNIV CALIFORNIA.

XX
PI Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;

XX
WPI; 1998-110587/10.

XX
PT New sequences from the 20q13 amplicon - used for detecting chromosomal abnormalities, particularly tumours, and for developing products for treating diseases.

XX
PS Claim 1; Page 62; 91pp; English.

XX
CC The sequence is that of a cDNA sequence cc43, which was isolated from the 20q13 amplicon. It is expressed in normal tissues but not been found in the breast cancer cell line. It can be used as a probe for the detection of chromosomal abnormalities at 20q13. It and other sequences isolated

from the 20q13 amplicon are consistently amplified in primary tumours. These sequences are useful as probes or as probe targets for monitoring the relative copy number of corresponding sequences from a biological sample such as tumour cells. The sequences can also be used in therapeutic applications for modulating the expression of the endogenous gene or the activity of the gene product. Examples of therapeutic approaches include antisense inhibition of gene expression, gene therapy and monoclonal antibodies that specifically bind the gene products. The products can also be used in the treatment of other diseases, e.g. age-related macular degeneration, Leber's congenital amaurosis and retinitis pigmentata

Seq	Sequence	2605 BP; 698 A; 625 C; 549 G; 732 T; 0 U; 1 Other;
	Query Match Best Local Similarity Matches 2604; Conservative	100.0%; Score 2604.6; DB 2; Length 2605; 100.0%; Pred. No. 0; 0; Mismatches 1; Indels 0; Gaps 0
Qy	1	CAAGCTCGAAATTAACCTCTACTAAAGGGAACAAAGCTGGAGCTCCACCGCGGTGGCGG 60
Db	1	CAAGCTCGAAATTAACCTCTACTAAAGGGAACAAAGCTGGAGCTCCACCGCGGTGGCGG 60
Qy	61	CCGCTCTAGAACTAGTGATCCCGCGGCTCAGAAATTCGGCAGAGCTGGCTACTTAC 120
Db	61	CCGCTCTAGAACTAGTGATCCCGCGGCTCAGAAATTCGGCAGAGCTGGCTACTTAC 120
Qy	121	GATGGCGATGATTTTCGAGTGGCCGTGGCAGTATCGCTTCCACCCCTTCTTTACGTTACA 180
Db	121	GATGGCGATGATTTTCGAGTGGCCGTGGCAGTATCGCTTCCACCCCTTCTTTACGTTACA 180
Qy	181	ACCGAATGTGACACTCGGCAGAAAGCACTGGCCCGCTGGTGTCTCGTGTGCTGTCTT 240
Db	181	ACCGAATGTGACACTCGGCAGAAAGCACTGGCCCGCTGGTGTCTCGTGTGCTGTCTT 240
Qy	241	CTGGCGCTGCACAAAAGTCCAGGATCAGCGTGAAGAGCTCAGGAGAGCCCGCTCTT 300
Db	241	CTGGCGCTGCACAAAAGTCCAGGATCAGCGTGAAGAGCTCAGGAGAGCCCGCTCTT 300
Qy	301	CAACAACTCAAGCTACAGCGAAAGCTTTCCTGTGGAGTCGATCCAGATGTGATTAGAGCA 360
Db	301	CAACAACTCAAGCTACAGCGAAAGCTTTCCTGTGGAGTCGATCCAGATGTGATTAGAGCA 360
Qy	361	ACTGAGGAAGAAAGGAAACCTCGAGTGGTTGGATAAGAGCAAGTCCAGCTTCTCGTATCAT 420
Db	361	ACTGAGGAAGAAAGGAAACCTCGAGTGGTTGGATAAGAGCAAGTCCAGCTTCTCGTATCAT 420
Qy	421	GTGGCGAGGCGCAGAAAGTATCGGGGAAACTCATCTATCAGTGGGTTCAGGAGTGGCCA 480
Db	421	GTGGCGAGGCGCAGAAAGTATCGGGGAAACTCATCTATCAGTGGGTTCAGGAGTGGCCA 480
Qy	481	GAACAACTCCGTCTTTTACCCCTGTATGAATCTGACTTAATCGGGAAGACACAGAGGATGAGCA 540
Db	481	GAACAACTCCGTCTTTTACCCCTGTATGAATCTGACTTAATCGGGAAGACACAGAGGATGAGCA 540
Qy	541	GTTTCCACGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGCAAGGCCCTTACAGCAGGAGCA 600
Db	541	GTTTCCACGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGCAAGGCCCTTACAGCAGGAGCA 600
Qy	601	CAAGGCCGAGATCATPACTGTACGCGATGGCGAGGCGTCAAGTTCTTCTAGCAGGAGACC 660
Db	601	CAAGGCCGAGATCATPACTGTACGCGATGGCGAGGCGTCAAGTTCTTCTAGCAGGAGACC 660
Qy	661	TGTTCTCCCTTTTACTTCTTTTACCTCCACCTTTTCAGGGCTTTTCAAAGAGACAGACCCAG 720
Db	661	TGTTCTCCCTTTTACTTCTTTTACCTCCACCTTTTCAGGGCTTTTCAAAGAGACAGACCCAG 720
Qy	721	TGTTCCCCCAAGACTGGATCTGTGACTCCACAGACTCAAAGGACTCCAGTCTCTGAAGG 780
Db	721	TGTTCCCCCAAGACTGGATCTGTGACTCCACAGACTCAAAGGACTCCAGTCTCTGAAGG 780
Qy	781	CTGGGACTGGGGATGGGTTTCTCACACCCCATATGTCTGTCCCTTGGATAGGGGTAGGC 840
Db	781	CTGGGACTGGGGATGGGTTTCTCACACCCCATATGTCTGTCCCTTGGATAGGGGTAGGC 840

Db 601 ACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTCCA 660
Qy 771 GTCCTGAAGCTGGAGACCTGGGGATGGTTCTTCACACCCATATGTCTGTCCTCCCTGGAT 830
Db 661 GTCCTGAAGCTGGAGACCTGGGGATGGTTCTTCACACCCATATGTCTGTCCTCCCTGGAT 720
Qy 831 AGGGTGAGGCTGAAGCACACAGGAGAAATATGTCTTCTTCGCGCTACCTCTCTTCC 890
Db 721 AGGGTGAGGCTGAAGCACACAGGAGAAATATGTCTTCTTCGCGCTACCTCTCTTCC 780
Qy 891 CATCTAGACTGCTCTGAGCCAGGCTCTGTAAACCTGCACACTTTATATGTGTTCCACACA 950
Db 781 CATCTAGACTGCTCTGAGCCAGGCTCTGTAAACCTGCACACTTTATATGTGTTCCACACA 840
Qy 951 TGTAAAGTACATACACATGCGCTGACGACATGCTTCTGCTCTCTCTCCACGCC 1010
Db 841 TGTAAAGTACATACACATGCGCTGACGACATGCTTCTGCTCTCTCTCCACGCC 900
Qy 1011 CTTTACGCTGCTTCCCTCTCTCAGGCTGCTGATCCTTCTAGGGGATGGGG 1070
Db 901 CTTTACGCTGCTTCCCTCTCTCAGGCTGCTGATCCTTCTAGGGGATGGGG 960
Qy 1071 AAGCCCTGGCTGAGGAGCCTTCCAGGCAATATGAAGATAGGAGGCCACCGGCTGGC 1130
Db 961 AAGCCCTGGCTGAGGAGCCTTCCAGGCAATATGAAGATAGGAGGCCACCGGCTGGC 1020
Qy 1131 AGTGAGAGGTGGGCCCCACACCGATTTATGATATTAATCTCAACTCCCAAAAAA 1190
Db 1021 AGTGAGAGGTGGGCCCCACACCGATTTATGATATTAATCTCAACTCCCAAAAAA 1080
Qy 1191 AAAAAAAA 1199
Db 1081 AAAAAAAA 1089

RESULT 3
AAI59037
ID AAI59037 standard; cDNA; 1243 BP.
XX
AC AAI59037;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1240.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR P-PSDB; AAM39881.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 1240; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1243 BP; 313 A; 321 C; 326 G; 283 T; 0 U; 0 Other;
Query Match 41.3%; Score 1076.2; DB 4; Length 1243;
Best Local Similarity 99.7%; Pred. No. 3e-211;
Matches 1078; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 113 GCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGAGTATCGTTCACACCTCTCTTT 172
Db 163 GCGTCGACGATGGCGATGAGTTTCGAGTGGCGGTGGAGTATCGTTCACACCTCTCTTT 222
Qy 173 ACCTTACAAACGATGTGGACACTCGGCAGAGCAGCTGGCGCTGGTGTGCTGGTGC 232
Db 223 ACCTTACAAACGATGTGGACACTCGGCAGAGCAGCTGGCGCTGGTGTGCTGGTGC 282
Qy 233 CTGTCTCTCTGCGCTGCACAAACAGTCCAGATGACGCGTGAAGAGTCCAGGAGAGC 292
Db 283 CTGTCTCTCTGCGCTGCACAAACAGTCCAGATGACGCGTGAAGAGTCCAGGAGAGC 342
Qy 293 CCGCTCTTCAACACAGTCAAGCTACAGGAAAGCTTCCTGTGAGTCCATCAGATTTGA 352
Db 343 CCGCTCTTCAACACAGTCAAGCTACAGGAAAGCTTCCTGTGAGTCCATCAGATTTGA 402
Qy 353 TTAGAGGAACCTGAGGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAGCTTC 412
Db 403 TTAGAGGAACCTGAGGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAGCTTC 462
Qy 413 CTGATCATGTGCGGAGGCCAGAAAGTGGGGAAACTCATCTATCAGTGGGTTTCCAGG 472
Db 463 CTGATCATGTGCGGAGGCCAGAAAGTGGGGAAACTCATCTATCAGTGGGTTTCCAGG 522
Qy 473 AGTGCCAGAAACACTCTCGTCTTTTACCTGTATGAACTGAATAATGGGGAAGACAGAG 532
Db 523 AGTGCCAGAAACACTCTCGTCTTTTACCTGTATGAACTGAATAATGGGGAAGACAGAG 582
Qy 533 GATGAGGATTTCCACGGGCTGGATGAAGCCACTCTACTGCGGCTCTGAGGCCCTACAG 592
Db 583 GATGAGGATTTCCACGGGCTGGATGAAGCCACTCTACTGCGGCTCTGAGGCCCTACAG 642
Qy 593 CAGGAGCAAAAGGCCGAGATCATCTGTTCAGCGATGGCCGAGGCGTCAAGTCTTCTAG 652
Db 643 CAGGAGCAAAAGGCCGAGATCATCTGTTCAGCGATGGCCGAGGCGTCAAGTCTTCTAG 702
Qy 653 CAGGAGCACTGTCTCTCTTTACTTTTACTCTCCACCTTTCCAGGCTTTCAAAAGAGAC 712
Db 703 CAGGAGCACTGTCTCTCTTTACTTTTACTCTCCACCTTTTCCAGGCTTTTCAAAAGAGAC 762

QY 1073 GCCTGGCTGCAGGCGCCTTCAGGCAATATGAGATAGGAGGCCAGCGGCTGGCAG 1132
 Db 1123 GCCTGGCTGCAGGCGCCTTCAGGCAATATGAGATAGGAGGCCAGCGGCTGGCAG 1182
 QY 1133 TGAGAGGTGTGGCCGCCACACCGGATTTATGATATTAAATCTCAACTCCCAAAAAA 1192
 Db 1183 TGAGAGGTGTGGCCGCCACACCGGATTTATGATATTAAATCTCAACTCCCAAAAAA 1242
 QY 1193 A 1193
 Db 1243 A 1243

RESULT 6
 AAI60823/c
 ID AAI60823 standard; cDNA; 1098 BP.
 AC AAI60823;
 XX
 DT 22-OCT-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 4812.
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.
 OS Homo sapiens.
 PN WO200153312-A1.
 PD 26-JUL-2001.
 PF 26-DEC-2000; 2000WO-US034263.
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR P-PSDB; AAM41667.
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 as central nervous system injuries.
 PS Claim 1; SEQ ID NO 4812; 10078pp; English.
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 encoded polypeptides (AAM38642-AAM42213) with nootropic,
 immunosuppressant and cytosstatic activity. The polynucleotides are useful
 in gene therapy. A composition containing a polypeptide or polynucleotide
 of the invention may be used to treat diseases of the peripheral nervous
 system, such as peripheral nervous injuries, peripheral neuropathy and
 localised neuropathies and central nervous system diseases, such as
 Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 utilisation of the activities such as: immune system suppression,
 Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 1098 BP; 245 A; 287 C; 308 G; 258 T; 0 U; 0 Other;
 Query Match 41.3%; Score 1075; DB 4; Length 1098;
 Best Local Similarity 99.9%; Pred. No. 5,1e-211;
 Matches 1086; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 108 GCTGGGCTACTAGGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCCCT 167
 Db 1087 GCTGGGCTACTAGGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCCCT 1028
 QY 168 TCTTTACGTTTCAACCGAATGTGGACACTCGGCAGAACAGCTGGCGCCTCGTGTCTCGC 227
 Db 1027 TCTTTACGTTTCAACCGAATGTGGACACTCGGCAGAACAGCTGGCGCCTCGTGTCTCGC 968
 QY 228 TGGTCTGTCTTCTGCGCGCTGCACAAACAGTCCAGCATGACGGTGTAGGAAGCTCAGG 287
 Db 967 TGGTCTGTCTTCTGCGCGCTGCACAAACAGTCCAGCATGACGGTGTAGGAAGCTCAGG 908
 QY 288 AGAGCCCGCTCTTCAACAGCTCAAGCTACAGCGAAAGCTTCTGTGGAGTCCATCCAGA 347
 Db 907 AGAGCCCGCTCTTCAACAGCTCAAGCTACAGCGAAAGCTTCTGTGGAGTCCATCCAGA 848
 QY 348 TTGTATTACAGAACTCGAGGAGAAAGGAACTCGAGTGGTGGATAGAGCAAGTCCA 407
 Db 847 TTGTATTACAGAACTCGAGGAGAAAGGAACTCGAGTGGTGGATAGAGCAAGTCCA 788
 QY 408 GCTTCTGATCATGTGGCGGAGCCAGAAAGATGGGGGAAACTCATCTATCAGTGGGTTT 467
 Db 787 GCTTCTGATCATGTGGCGGAGCCAGAAAGATGGGGGAAACTCATCTATCAGTGGGTTT 728
 QY 468 CCAGGATGGCCAGAACAACTCGGTCTTTACCTGTATGAACCTGACTTAATGGGAGACA 527
 Db 727 CCAGGATGGCCAGAACAACTCGGTCTTTACCTGTATGAACCTGACTTAATGGGAGACA 668
 QY 528 CAGAGGATGAGGATTTCCACGGCTGGATGAAGCCACTCTACTCGGGGCTCTCGAGGCC 587
 Db 667 CAGAGGATGAGGATTTCCACGGCTGGATGAAGCCACTCTACTCGGGGCTCTCGAGGCC 608
 QY 588 TACAGCAGGACCAAGGCCGAGATCATCACTGTTCAGCGATGG - CCGAGCGCTCAAGTTC 646
 Db 607 TACAGCAGGACCAAGGCCGAGATCATCACTGTTCAGCGATGGCCCGAGSGCTCAAGTTC 548
 QY 647 TTCTAGCAGGACCTGTCTCCCTTTACTTTACTTCTTACCTCCACCTTTCCAGGGCTTCAAAA 706
 Db 547 TTCTAGCAGGACCTGTCTCCCTTTACTTTACTTCTTACCTCCACCTTTCCAGGGCTTCAAAA 488
 QY 707 GGAGACAGACCCAGGTGTCCCAAGACCTGGATCTGTGACTCCACAGACTCAAAAGGAC 766
 Db 487 GGAGACAGACCCAGGTGTCCCAAGACCTGGATCTGTGACTCCACAGACTCAAAAGGAC 428
 QY 767 TCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGTCCCTT 826
 Db 427 TCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGTCCCTT 368
 QY 827 GGATAGGTGAGCTGAGACACACAGGGAGAAATATGTGCTTCTTCTCGCCCTACCTCCT 886
 Db 367 GGATAGGTGAGCTGAGACACACAGGGAGAAATATGTGCTTCTTCTCGCCCTACCTCCT 308
 QY 887 TTCCCATCTAGACTGTCTCTTGGCCAGGGTCTGTAAACCTGACACTTTATATGTGTCA 946
 Db 307 TTCCCATCTAGACTGTCTCTTGGCCAGGGTCTGTAAACCTGACACTTTATATGTGTCA 248
 QY 947 CACATGTAGTACATACACATATGGCCCTGACGACATGCTTCTGTCTCTCTCTCTCC 1006
 Db 247 CACATGTAGTACATACACATATGGCCCTGACGACATGCTTCTGTCTCTCTCTCTCC 188
 QY 1007 ACCCTTTAGTGTGTGCTTGGCTTCTTCTGAGGCTGGTGTGATCCTTCTTAGGGGATG 1066

Db 187 ACCCTTTAGCTGCTGTTGGCTCCCTTCTCAGGCTGGTGTGATCTCTTCAGGGATG 128

QY 1067 GGGGAGCCCTGGCTGCGAGGAGCCCTTCAGGCAATATAGATAGAGGCCCGGCC 1126

Db 127 GGGGAAGCCCTGGCTGCGAGGAGCCCTTCAGGCAATATAGATAGAGGCCCGGCC 68

QY 1127 TGGCAGTGCAGAGGTGTGGCCGCCACACCGGATTTATGATATTAATACTCAACTCCCAAAA 1186

Db 67 TGGCAGTGCAGAGGTGTGGCCGCCACACCGGATTTATGATATTAATACTCAACTCCCAAAA 8

QY 1187 AAAAAA 1193

Db 7 AAAAAA 1

RESULT 7

AAH68547

ID AAH68547 standard; cDNA; 1087 BP.

AC AAH68547;

XX 13-SEP-2001 (first entry)

DT Human protein HP10149 coding sequence.

DE Human; gene therapy; tumour; ss.

XX Homo sapiens.

OS WO200142302-A1.

FN 14-JUN-2001.

XX 06-DEC-2000; 2000WO-JP008631.

PF 06-DEC-1999; 99JP-00346863.

PR 06-DEC-1999; 99JP-00346864.

PR 08-FEB-2000; 2000JP-00031062.

PR 10-FEB-2000; 2000JP-00034090.

PR 10-FEB-2000; 2000JP-00034091.

PR 14-FEB-2000; 2000JP-00035829.

PR 14-FEB-2000; 2000JP-00035899.

PR 14-MAR-2000; 2000JP-00071161.

PR 30-MAY-2000; 2000JP-00160851.

XX (NTSC-) JAPAN SCI & TECHNOLOGY CORP.

PA Kato S, Eguchi C, Saeki M;

FI WPI; 2001-381646/40.

DR P-PSDB; AAG93262.

XX Human protein originated from tumor cell line, applicable as drug,

PT reagent for studying intracellular protein networks and protein source

PT for drug screening, also encoded cDNA for gene diagnosis and gene

PT therapy.

XX Claim 3; Page 169-171; 471pp; Japanese.

PS The present sequence is a human protein coding sequence. The human

XX protein, preferably originated from tumour cell line, is applicable as a

CC drug, a reagent for studying intracellular protein networks and a protein

CC source for screening proteins for binding low molecular weight drugs. The

CC human protein coding sequence is useful for gene diagnosis and gene

CC therapy, expression vectors and transformant cells for detection of

XX ligands and receptors

Sequence 1087 BP; 246 A; 304 C; 288 G; 249 T; 0 U; 0 Other;

Query Match 41.2%; Score 1073; DB 5; Length 1087;

Best Local Similarity 100.0%; Pred. No. 1.3e-210;

Matches 1073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 CTGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTCCCAACCTT 168

Db 15 CTGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTCCCAACCTT 74

QY 169 CTTTACGTTTAAACCGAATGTGGACACTCGGCAGAACAGCTGGCCGCTGGTGTCTGCT 228

Db 75 CTTTACGTTTAAACCGAATGTGGACACTCGGCAGAACAGCTGGCCGCTGGTGTCTGCT 134

QY 229 GGTCTGTCTTCTGCGCCCTGCCAACAACAGTCCAGCATGCGGTGATGGAGTCCAGGA 288

Db 135 GGTCTGTCTTCTGCGCCCTGCCAACAACAGTCCAGCATGCGGTGATGGAGTCCAGGA 194

QY 289 GAGCCCGCTCTTCAACAACGTCAGCTACAGCGAAGCTTCTGTGGAGTGCATCCAGAT 348

Db 195 GAGCCCGCTCTTCAACAACGTCAGCTACAGCGAAGCTTCTGTGGAGTGCATCCAGAT 254

QY 349 TGTATTAGAGAACTGAGGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAG 408

Db 255 TGTATTAGAGAACTGAGGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAG 314

QY 409 CTTCTGTATCATGTGGCGGAGCCAGAGAAATGGGGAAACTCATCTATCATGTGGGTTTC 468

Db 315 CTTCTGTATCATGTGGCGGAGCCAGAGAAATGGGGAAACTCATCTATCATGTGGGTTTC 374

QY 469 CAGGAGTGGCCAGAACTCCGCTCTTTACCTGTATGAACCTGACTAATGGGGAAGACAC 528

Db 375 CAGGAGTGGCCAGAACTCCGCTCTTTACCTGTATGAACCTGACTAATGGGGAAGACAC 434

QY 529 AGAGATGAGAGTTTCCACGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGCGAGCCCT 588

Db 435 AGAGATGAGAGTTTCCACGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGCGAGCCCT 494

QY 589 ACAGCAGGAGCACAGGCGGAGATCATCACTGTGAGCGATGGCGAGGCGTCAAGTCTT 648

Db 495 ACAGCAGGAGCACAGGCGGAGATCATCACTGTGAGCGATGGCGAGGCGTCAAGTCTT 554

QY 649 CTAGCAGGAGCTGTCTCCCTTTACTTCTTACTCTCCACCTTTCAGGGCTTTCAAAAGG 708

Db 555 CTAGCAGGAGCTGTCTCCCTTTACTTCTTACTCTCCACCTTTCAGGGCTTTCAAAAGG 614

QY 709 AGACAGACCCAGTGTCCCCAAAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTC 768

Db 615 AGACAGACCCAGTGTCCCCAAAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTC 674

QY 769 CAGTCTGAAGCTGGGACCTGGGGATGGGTTCTCACACCCCATATCTGTCTCTCTGG 828

Db 675 CAGTCTGAAGCTGGGACCTGGGGATGGGTTCTCACACCCCATATCTGTCTCTCTGG 734

QY 829 ATAGGCTGAGGCTGAAGCACAGGAGAAATAATGTGCTTCTTCTCGCCCTACCTCTTT 888

Db 735 ATAGGCTGAGGCTGAAGCACAGGAGAAATAATGTGCTTCTTCTCGCCCTACCTCTTT 794

QY 889 CCCATCTAGACTGTCTTGAGCCAGGGTCTGTAAACCTGACACTTATATGTGTTTACA 948

Db 795 CCCATCTAGACTGTCTTGAGCCAGGGTCTGTAAACCTGACACTTATATGTGTTTACA 854

QY 949 CATGTAAGTACATACACATGCGCTCGACACATGCTTCTGTCTCTCTCTCTCCAC 1008

Db 855 CATGTAAGTACATACACATGCGCTCGACACATGCTTCTGTCTCTCTCTCTCTCCAC 914

QY 1009 CCCTTTAGCTGTGTGCTCTCCCTTCTCAGGCTGGTGTGGATCTCTCTAGGGATGGG 1068

Db 915 CCCTTTAGCTGTGTGCTCTCCCTTCTCAGGCTGGTGTGGATCTCTCTAGGGATGGG 974

QY 1069 GGAAGCCCTGGCTGAGGAGCGCTTTCAGGCAATATGAAGATAGGAGGCCACAGGCGCTG 1128

Db 975 GGAAGCCCTGGCTGAGGAGCGCTTTCAGGCAATATGAAGATAGGAGGCCACAGGCGCTG 1034

QY 1129 GCAGTGAGAGGTGTGGCCCCCACACCGATTTATGATATTAATAATCTCAACTCCC 1181

Db 1035 GCAGTGAGAGGTGTGGCCCCCACACCGATTTATGATATTAATAATCTCAACTCCC 1087

RESULT 8
ADQ84631
ID ADQ84631 standard; cDNA; 1100 BP.
XX
AC ADQ84631;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1445.
XX
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
XX
PA (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
PI Wu TD, Zhou Y;
XX
XX WPI; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 1445; 5504pp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)
XX a process for producing a polypeptide; (4) an isolated polypeptide
XX comprising: (a) an amino acid sequence encoded by any of the above
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-
XX length coding region of the above nucleotide sequences; or (c) a sequence
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)
XX an isolated antibody that binds to the above polypeptide; (7) a process
XX for producing the antibody; (8) an isolated oligopeptide that binds to
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)
XX binding organic molecule that binds to the above polypeptide; (10) a
XX composition of matter comprising the above (chimeric) polypeptide,
XX antibody, oligopeptide or TAT binding organic molecule, in combination
XX with a carrier; (11) an article of manufacture comprising a container and
XX the composition of matter contained within the container; (12) methods of
XX inhibiting the growth of a cell that expresses the above protein, where
XX the growth of the cell is at least in part dependent upon a growth
XX potentiating effect of the above protein; (13) a method of
XX therapeutically treating a mammal having a cancerous tumour comprising
XX cells that express the above protein; (14) a method of determining the
XX presence of a protein in a sample suspected of containing the protein
XX described above; (15) methods of diagnosing the presence of a tumour in a
XX mammal; (16) a method for treating or preventing a cell proliferative
XX disorder associated with increased expression or activity of the above
XX protein; and (17) a method of binding an antibody, oligopeptide or
XX organic molecule to a cell that expresses the protein described above.
XX The TAT sequences have cytostatic activities, and can be used in gene
XX therapy. The composition and methods are useful for diagnosing,
XX preventing or treating cancer. The composition is also used for preparing
XX a medicament for the therapeutic treatment or diagnostic detection of a
XX cell proliferative disorder or cancer. The present sequence represents a

CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 1100 BP; 247 A; 308 C; 292 G; 253 T; 0 U; 0 Other;
Query Match 41.2%; Score 1073; DB 13; Length 1100;
Best Local Similarity 100.0%; Pred. No. 1.3e-210;
Matches 1073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 109 CTGGGCTACTACGATGGCGATGAGTTTCAGTGGCGTGGCAGATATCGCTTCCACCCCTT 168
DB 28 CTGGGCTACTACGATGGCGATGAGTTTCAGTGGCGTGGCAGATATCGCTTCCACCCCTT 87
QY 169 CTTTACGTTTAAACCGAATGTGGACACTCGGCAGAAAGAGCTGGCCCTCTGTGCTCGCT 228
DB 88 CTTTACGTTTAAACCGAATGTGGACACTCGGCAGAAAGAGCTGGCCCTCTGTGCTCGCT 147
QY 229 GGTCTGTCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTATGGAAGCTCAGGA 288
DB 148 GGTCTGTCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTATGGAAGCTCAGGA 207
QY 289 GAGCCGCTCTTCAACCAACGCTCAAGCTACAGCGAAAGCTTCTCTGTGGAGTCCGATCCAGAT 348
DB 208 GAGCCGCTCTTCAACCAACGCTCAAGCTACAGCGAAAGCTTCTCTGTGGAGTCCGATCCAGAT 267
QY 349 TGTATTAGAGGAACTGAGGAAAGAGGAACTCTCGAGTGGTGGATAGAGCAAGTCCAG 408
DB 268 TGTATTAGAGGAACTGAGGAAAGAGGAACTCTCGAGTGGTGGATAGAGCAAGTCCAG 327
QY 409 CTTCTGTATCATGTGGCGGAGGCCAGAGAAATGGGGGAACTCATCTATCAGTGGGTTTC 468
DB 328 CTTCTGTATCATGTGGCGGAGGCCAGAGAAATGGGGGAACTCATCTATCAGTGGGTTTC 387
QY 469 CAGGAGTGCCAGAACTCCGCTCTTTACCTCTATGAACTGACTTAATGGGGAAGACAC 528
DB 388 CAGGAGTGCCAGAACTCCGCTCTTTACCTCTATGAACTGACTTAATGGGGAAGACAC 447
QY 529 AGAGGATGAGGAGTTCACGGGCTGGATGAAGCACTTACTTCCAGGCTTTCAGAGCCCT 588
DB 448 AGAGGATGAGGAGTTCACGGGCTGGATGAAGCACTTACTTCCAGGCTTTCAGAGCCCT 507
QY 589 ACAGCAGGACACAGGCGGAGATCATCTCTCAGCGATGCGGAGGCGTCAAGTTCCT 648
DB 508 ACAGCAGGACACAGGCGGAGATCATCTCTCAGCGATGCGGAGGCGTCAAGTTCCT 567
QY 649 CTAGCAGGACCTGTCTCTCTTACTTCTTACTTCCACCTTTCAGAGGCTTTCAGAGG 708
DB 568 CTAGCAGGACCTGTCTCTCTTACTTCTTACTTCCACCTTTCAGAGGCTTTCAGAGG 627
QY 709 AGACAGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTC 768
DB 628 AGACAGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTC 687
QY 769 CAGTCTGAGGCTGGGACCTGGGATGGGTTTCTCACACCCCATATGTCTGCTCCCTGG 828
DB 688 CAGTCTGAGGCTGGGACCTGGGATGGGTTTCTCACACCCCATATGTCTGCTCCCTGG 747
QY 829 ATAGGCTGAGGCTGAAGCAGCAGGAGAAATATGTGCTTCTTCTCGCCCTACTCTCTT 888
DB 748 ATAGGCTGAGGCTGAAGCAGCAGGAGAAATATGTGCTTCTTCTCGCCCTACTCTCTT 807
QY 889 CCCATCTAGACTGTCTTGGACCCAGGGTCTGTAAACCTGACACTTATATGTGTTTACA 948
DB 808 CCCATCTAGACTGTCTTGGACCCAGGGTCTGTAAACCTGACACTTATATGTGTTTACA 867
QY 949 CATGTAAGTACATACACATGCGGCTGAGCAGACATGCTTCTGCTCTCTCTCTCCAC 1008
DB 868 CATGTAAGTACATACACATGCGGCTGAGCAGACATGCTTCTGCTCTCTCTCTCCAC 927
QY 1009 CCCTTTAGTGTGTTGCTCTCTCTCAGGCTGTGTGATCTCTCTCAGGAGATGGG 1068
DB 928 CCCTTTAGTGTGTTGCTCTCTCTCAGGCTGTGTGATCTCTCTCAGGAGATGGG 987
QY 1069 GGAAGCCCTGGCTGCAGGCGAGCCTTTCAGGCAATATGAGATAGGAGGCCCGGCGCTG 1128

```

Db      988  GGAACCCCTGGCTGCAGCAGCCTTCCAGGCAATATAGATAGGAGGCCCAAGGCGCTG 1047
QY      1129  GCAGTGAGAGGTGTGGCCCCCACCAGATTATGATATTAAATCTCAACTCCC 1181
Db      1048  GCAGTGAGAGGTGTGGCCCCCACCAGATTATGATATTAAATCTCAACTCCC 1100

RESULT 9
ABQ57032/c
ID      ABQ57032 standard; cDNA; 634 BP.
XX      AC      ABQ57032;
XX      XX
XX      DT      02-AUG-2002 (first entry)
XX      DE      Human colon cancer related nucleotide sequence SEQ ID NO:727.
XX      KW      Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX      KW      genetic analysis; diagnostic; antisense therapy; gene; ss.
XX      OS      Homo sapiens.
XX      PN      WO200229086-A2.
XX      PD      11-APR-2002.
XX      PF      02-OCT-2001; 2001WO-US030732.
XX      PR      02-OCT-2000; 2000US-0237271P.
XX      PA      (FARB ) BAYER CORP.
XX      PI      Burgess C, Astle JH, Carroll E, Catino TU, Dwivedi P, Molino GA;
XX      PI      Thiallingam A, Lewis MB;
XX      DR      WPI; 2002-426115/45.
XX      PT      New isolated nucleic acid that is differentially expressed in cancer
XX      PT      tissues useful for determining the presence of colon cancer in a cell or
XX      PT      tissue type, and in antisense therapy.
XX      PS      Claim 1; Fig 1; 796pp; English.
XX      CC      ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
XX      CC      expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
XX      CC      encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
XX      CC      used in antisense therapy. An antibody immunoreactive with a polypeptide
XX      CC      encoded by (I) is useful for detecting cancer in a patient sample, and
XX      CC      for detecting the presence or absence of a polynucleotide encoded by a
XX      CC      nucleic acid which hybridizes to (I) in a cell. A probe/primer derived
XX      CC      from (I) can be used for determining the presence of a nucleic acid which
XX      CC      hybridizes to (I), and for determining the phenotype of cells in a sample
XX      CC      of cells from a patient. (I) is useful for determining the presence of
XX      CC      colon cancer in a cell or tissue type, for determining the presence or
XX      CC      state of other type of cancer, in antisense therapy, to generate
XX      CC      macroarrays on a solid surface, to identify a chromosome on which the
XX      CC      corresponding gene resides, and in tissue profiling, forensics, genetic
XX      CC      analysis, mapping and diagnostic applications. (I) can be used to raise
XX      CC      antibodies, and to screen for peptide analogues and antagonists
XX      SQ      Sequence 634 BP; 147 A; 168 C; 160 G; 158 T; 0 U; 1 Other;

Query Match      23.2%; Score 603.6; DB 6; Length 634;
Best Local Similarity 98.7%; Pred. No. 3.5e-114;
Matches 629; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY      321  GAAGCTTCTCTGAGTCGATCCAGATTGTATTAGAGCACTGAGGAAGGGAACC 380
Db      634  GAAGCTTCTCTGAGTCGATCCAGATTGTATTAGAGAA--TTAGAGAAAGGGAACN 577
QY      381  TCGAGTGGTTGGATAAGAGCAAGTCAGCTTCTCTGATCATGTGGCGGAGCCAGAGAAT 440

```

```

Db      576  TCGATGGTTGGATAAGAGCAAGTCCAGCTTCTCTGATCATGTGGCGGAGCCAGAGAAT 517
QY      441  GGGGGAATCTCATCTATCAGTGGGTTTCCAGAGGTGGCCAGAACAACTCCGTCCTTTACCC 500
Db      516  GGGGGAATCTCATCTATCAGTGGGTTTCCAGAGGTGG--CAGAACAACTCCGTCCTTTACCC 458
QY      501  TGTATGAATCTACTAATGGGGAAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAG 560
Db      457  TGTATGAATCTACTAATGGGGAAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAG 398
QY      561  CCACTTACTCGGGCTCTGCAGGCCCTTACAGCAGGAGCAAGGCCGAGATCATCATCTG 620
Db      397  CCACTTACTCGGGCTCTGCAGGCCCTTACAGCAGGAGCAAGGCCGAGATCATCATCTG 338
QY      621  TCAGCGATGGCGAGGCGTCAAGTTCTTCTAGCAGGAGCTGTCTCCCTTTACTTCTTAC 680
Db      337  TCAGCGATGGCGAGGCGTCAAGTTCTTCTAGCAGGAGCTGTCTCCCTTTACTTCTTAC 278
QY      681  CTCCACCTTTCCAGGGCTTTCAAAAGGAGACAGACCAGTGTCCCCCAAGACTGGATC 740
Db      277  CTCCACCTTTCCAGGGCTTTCAAAAGGAGACAGACCAGTGTCCCCCAAGACTGGATC 218
QY      741  TGTGACTCCACAGACTCAAAAGGACTCCAGTCTCTGAAGGCTGGACCTGGGGATGGGTT 800
Db      217  TGTGACTCCACAGACTCAAAAGGACTCCAGTCTCTGAAGGCTGGACCTGGGGATGGGTT 158
QY      801  TCTCACACCCATATGTCTGTCCCTTGGATAGGCTGAGGCTGAAGCACAGGAGAAAT 860
Db      157  TCTCACACCCATATGTCTGTCCCTTGGATAGGCTGAGGCTGAAGCACAGGAGAAAT 98
QY      861  ATGTGCTTCTTCTCGCCCTACCTCTCTTCCATCTAGACTGTCTCTTGAGCCAGGCTCTG 920
Db      97  ATGTGCTTCTTCTCGCCCTACCTCTCTTCCATCTAGACTGTCTCTTGAGCCAGGCTCTG 38
QY      921  TAAACCTGACACTTATATGTGTTTCACACATGTAAAT 957
Db      37  TAAACCTGGCACTTATATGTGTTTCACACATGTAAAT 1

RESULT 10
ABQ56941
ID      ABQ56941 standard; cDNA; 559 BP.
XX      AC      ABQ56941;
XX      DT      02-AUG-2002 (first entry)
XX      DE      Human colon cancer related nucleotide sequence SEQ ID NO:636.
XX      KW      Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX      KW      genetic analysis; diagnostic; antisense therapy; gene; ss.
XX      OS      Homo sapiens.
XX      PN      WO200229086-A2.
XX      PD      11-APR-2002.
XX      PF      02-OCT-2001; 2001WO-US030732.
XX      PR      02-OCT-2000; 2000US-0237271P.
XX      PA      (FARB ) BAYER CORP.
XX      PI      Burgess C, Astle JH, Carroll E, Catino TU, Dwivedi P, Molino GA;
XX      PI      Thiallingam A, Lewis MB;
XX      DR      WPI; 2002-426115/45.
XX      PT      New isolated nucleic acid that is differentially expressed in cancer
XX      PT      tissues useful for determining the presence of colon cancer in a cell or
XX      PT      tissue type, and in antisense therapy.

```


PS	Claim 1; Fig 1; 796pp; English.	
XX	ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be used in antisense therapy. An antibody immunoreactive with a polypeptide encoded by (I) is useful for detecting cancer in a patient sample, and for detecting the presence or absence of a polynucleotide encoded by a nucleic acid which hybridizes to (I) in a cell. A probe/primer derived from (I) can be used for determining the presence of a nucleic acid which hybridizes to (I), and for determining the phenotype of cells in a sample of cells from a patient. (I) is useful for determining the presence of colon cancer in a cell or tissue type, for determining the presence or state of other type of cancer, in antisense therapy, to generate macroarrays on a solid surface, to identify a chromosome on which the corresponding gene resides, and in tissue profiling, forensics, genetic analysis, mapping and diagnostic applications. (I) can be used to raise antibodies, and to screen for peptide analogues and antagonists	
XX		
XX	Sequence 559 BP; 135 A; 140 C; 167 G; 117 T; 0 U; 0 Other;	
XX		
XX	Query Match 19.8%; Score 515.8; DB 6; Length 559;	
XX	Best Local Similarity 99.6%; Pred. No. 3.8e-96;	
XX	Matches 517; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	109 CTGGGCTACTAGATGGCGATGATTTTCAGTGGCGCGTGGCAGATATCGCTTCCACCCCTT 168	
DB	16 CTGGGCTACTAGATGGCGATGATTTTCAGTGGCGCGTGGCAGATATCGCTTCCACCCCTT 75	
QY	169 CTTTAGGTTACACCAATGTGGACACTCGGCAGAGCAGCTGGCGCGTGGTGTCTCGCT 228	
DB	76 CTTTAGGTTACACCAATGTGGACACTCGGCAGAGCAGCTGGCGCGTGGTGTCTCGCT 135	
QY	229 GGTCTCTGCTCTTCTGCGCGCTGCACAAACAGTCCAGCATGACGGTGTATGGAAGCTCAGGA 288	
DB	136 GGTCTCTGCTCTTCTGCGCGCTGCACAAACAGTCCAGCATGACGGTGTATGGAAGCTCAGGA 195	
QY	289 GAGCCCGCTCTTCAACAAAGTCAAGCTACAGCAAGAAAGCTTCTCTGTGGATGTCAGATCCAGAT 348	
DB	196 GAGCCCGCTCTTCAACAAAGTCAAGCTACAGCAAGAAAGCTTCTCTGTGGATGTCAGATCCAGAT 255	
QY	349 TGTATTAGAGGACTGAGGAGAGAGGAAAGCCTCGAGTGGTGGATAGAGCAAGTCCAG 408	
DB	256 TGTATTAGAGGAACTGAGGAGAGAGGAAAGCCTCGAGTGGTGGATAGAGCAAGTCCAG 315	
QY	409 CTTCTCTGATCTGTGGCGAGGCCAGAAAGTGGGGAAACTCATCTATCAGTGGGTTTC 468	
DB	316 CTTCTCTGATCTGTGGCGAGGCCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTTC 375	
QY	469 CAGGAGTGCCCGAGAACTCCGTCTTTTACCCTGTATGAACCTGACTTAATGGGGAAGACAC 528	
DB	376 CAGGAGTGCCCGAGAACTCCGTCTTTTACCCTGTATGAACCTGACTTAATGGGGAAGACAC 435	
QY	529 AGAGGATAGGAGTTCCAGGGCTGATGAGCCACTCTACTCGGGCTCTGCAGGCCCT 588	
DB	436 AGAGGATAGGAGTTCCAGGGCTGATGAGCCACTCTACTCGGGCTCTGCAGGCCCT 495	
QY	589 ACAGCAGGAGCAAGCGCGAGATCATCTGTCCAGCA 627	
DB	496 ACAGCAGGAGCAAGCGCGAGATCATCTGTCCAGCA 534	
DE	RESULT 11	
XX	ADQ50360	
XX	ID ADQ50360 standard; DNA; 675 BP.	
XX	AC ADQ50360;	
XX	21-OCT-2004 (first entry)	
DE	Novel canine microarray-related DNA sequence SeqID1662.	
XX	canine microarray; drug screening; toxicity assay;	
XX		
XX	environmental pollutant; cellular response; gene expression profile; toxic response; liver necrosis; fatty liver disease; protein adduct formation; hepatitis; dog; ds.	
XX	Canis familiaris.	
XX	WO2004063324-A2.	
XX	29-JUL-2004.	
XX	05-MAY-2003; 2003WO-US013853.	
XX	03-MAY-2002; 2002US-0377240P.	
XX	(GENE-) GENE LOGIC INC.	
XX	(PFIZ) PFIZER PROD INC.	
XX	Diggans JC, Porter M, Wei T;	
XX	WPI; 2004-561890/54.	
XX	New isolated nucleic acid molecule, useful for drug screening and toxicity assays or for assessing the impact, including toxicity, of a compound, pharmaceutical agent or environmental pollutant on a cell or living organism.	
XX	Claim 1; SEQ ID NO 1662; 41pp; English.	
XX	This invention is related to a novel isolated canine nucleic acid sequences and the construction of canine microarrays containing a significant portion of the canine genome. The isolated canine nucleic acid sequences of the invention may be useful for drug screening and toxicity assays. The invention is therefore useful for assessing the impact, including toxicity, of a compound, pharmaceutical agent or environmental pollutant on a cell or living organism. The methods are useful for detecting genes that are up- or down-regulated in canines in a disease state. The sequences are useful as diagnostic agents or markers to detect a cellular response in a sample individually or as part of a gene expression profile. It is also useful as a target for agents that modulate gene expression or activity. The database is useful for producing electronic Northern that allow the user to determine the cell type or tissue in which a given gene is expressed and to allow determination of the abundance or expression level of a given gene in a particular tissue or cell. The methods are useful for determining the similarity of a toxic response to one or more individual compounds. The methods are useful for predicting at least one toxic response or the likelihood that a compound or test agent will induce various specific pathologies such as those of the liver (liver necrosis, fatty liver disease, protein adduct formation or hepatitis), those of the kidney, heart, brain or testes, or other pathologies associated with at least one of the toxins. The methods are also useful for predicting or elucidating the potential cellular pathways influenced, induced or modulated by the compound or test agent due to the similarity of the expression profile compared to the profile induced by a known toxin. The present sequence is that of a canine DNA sequence which was claimed for use during the production of a canine microarray of the invention.	
XX	Sequence 675 BP; 159 A; 189 C; 170 G; 155 T; 0 U; 2 Other;	
XX		
XX	Query Match 18.8%; Score 489.2; DB 13; Length 675;	
XX	Best Local Similarity 86.2%; Pred. No. 1.2e-90;	
XX	Matches 595; Conservative 0; Mismatches 70; Indels 25; Gaps 4;	
QY	433 AGAAGAATGGGGAACTCATCTATCAGTGGGTTTCCAGAGTGGCCAGAACAACTCCGT 492	
DB	1 AGAAGAATGGGGAACTCATCTATCAGTGGGTTTCTANNAGTGGCCAGAACAACTCCGT 60	
QY	493 CTTTACCCCTGTATGAATGACTAATGGGGAAGACACAGAGGATGAGAGTTCCACGGCT 552	
DB	61 GTTCACCCCTGTATGAATGACTAATGGGGAAGACACAGAGGATGAGAGTTCCATGGCT 120	
QY	553 GGATGAGCCACTCTACTCGGGCTCTGAGGCCCTTACAGCAGGACACAGGCCAGAT 612	

genetic analysis; diagnostic; antisense therapy; gene; ss.
Homo sapiens.
WO200229086-A2.
11-APR-2002.
02-OCT-2001; 2001WO-US030732.
02-OCT-2000; 2000US-0237271P.
(FARB) BAYER CORP.
Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA, Thiagalingam A, Lewis ME;
WPI; 2002-426115/45.
New isolated nucleic acid that is differentially expressed in cancer tissues useful for determining the presence of colon cancer in a cell or tissue type, and in antisense therapy.
Claim 1; Fig 1; 796pp; English.
ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially expressed in cancer tissues. ABB78993 to ABB79004 represent proteins encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be used in antisense therapy. An antibody immunoreactive with a polypeptide encoded by (I) is useful for detecting cancer in a patient sample, and for detecting the presence or absence of a polynucleotide encoded by a nucleic acid which hybridizes to (I) in a cell. A probe/primer derived from (I) can be used for determining the presence of a nucleic acid which hybridizes to (I), and for determining the phenotype of cells in a sample of cells from a patient. (I) is useful for determining the presence of colon cancer in a cell or tissue type, for determining the presence or state of other type of cancer, in antisense therapy, to generate macroarrays on a solid surface, to identify a chromosome on which the corresponding gene resides, and in tissue profiling, forensics, genetic analysis, mapping and diagnostic applications. (I) can be used to raise antibodies, and to screen for peptide analogues and antagonists
Sequence 609 BP; 156 A; 149 C; 160 G; 135 T; 0 U; 9 Other;
Query Match 15.2%; Score 395; DB 6; Length 609;
Best Local Similarity 96.5%; Pred. No. 2.6e-71;
Matches 447; Conservative 0; Mismatches 10; Indels 6; Gaps 4;
112 GGCTACTAGTGGCGATGAGTTTCGAGTGGCGGTGGCGATATCGCTTCCACCTTCTT 171
|||
5 GGGGACTACGATGGTGTGATGAGTTTCGAGTGGCGGTGGCGATATCGCTTCCACCTTCTT 64
172 TACGTTCACCAACGATGGACACTCGGCAGAGCAGCTGGCGCTGGTGGCTCGCTCGT 231
|||
65 TACGTTCACCAACGATGGACACTCGGCAGAGCAGCTGGCGCTGGTGGCTCGCTCGT 124
232 CCTGTCTCTTCTCCGCTGCACAAACAGTCCAGCATGACGGTGGTGGAGCTCAGAGAG 291
135 CCTGTCTCTTCTCCGCTGCACAAACAGTCCAGCATGACGGTGGTGGAGCTCAGAGAG 184
292 CCCGTCTTCAACAAACGTCACAGTACAGCAAGCTTCTGTGGAGTCGATCCAGATTGT 351
185 CCCGTCTTCAACAAACGTCACAGTACAGCAAGCTTCTGTGGAGTCGATCCAGATTGT 244
352 ATTAGGAGACTGAGGAGAAAGGAACTCCAGTGGTGGATAGAGCAAGTCCAGCTT 411
245 ATTAGGAGAACTGAGGAGAAAGGAACTCCAGTGGTGGATAGAGCAAGTCCAGCTT 304
412 CCTGTATCATGTGGCGAGCCAGAGAAATGGGGGAAACTCATCTATCATGTGGTTTCAG 471
305 CCTGTATCATGTGGCGGA-GCCAGAGAAATGGGGGAAACTCATCTATCATGTGG--TTTCAG 361
472 GAGTGCCCAAGAACTCCGCTTTTACCCCTGTATGAAGTACTGACTAATGGGGAGACACAGA 531

Db 362 GAGTGGC--AGAACACACTCGCTCTTTACCTCTGATGAACCTACTAATGGGAGACACAGA 419
Qy 532 GGATGAGGAGTTCCACGGGCTGGATGAAGCCACTCTTACTGCGG 574
Db 420 GGATGAGGAGTTCCACGGGCTGGATGAAG--CCACTTACTGCGG 461
RESULT 14
ABX17514
ID ABX17514 standard; cDNA; 545 BP.
XX AC ABX17514;
XX 05-FEB-2003 (first entry)
XX Human cDNA encoding protein NOV5.
XX Human; ss; gene; NOVX; adrenoleukodystrophy; haemophilia; stoke; VHL;
KW congenital adrenal hyperplasia; haemophilia; hypercoagulation;
KW idiopathic thrombocytopenic purpura; autoimmune disease; allergy;
KW immunodeficiencies; transplantation; Von Hippel-Lindau syndrome;
KW Alzheimer's disease; tuborous sclerosis; Parkinson's disease; epilepsy;
KW Huntington's disease; cerebral palsy; Lesch-Nyhan syndrome; pain;
KW multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety;
KW behavioural disorder; addiction; neuroprotection; diabetes; ARDS;
KW polycystic kidney disease; systemic lupus erythematosus; IGA;
KW renal tubular acidosis; immunoglobulin A nephropathy; hypercalcaemia;
KW cirrhosis; transplantation; asthma; emphysema; scleroderma; GWD;
KW adult respiratory distress syndrome; graft versus host disease;
KW lymphedema; fertility; pancreatitis; obesity; haemophilia; ulcer;
KW anaemia; cancer; trauma; regeneration; infection.
XX Homo sapiens.
OS WO200281629-A2.
XX 17-OCT-2002.
XX 03-APR-2002; 2002WO-US010522.
XX 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 05-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 19-APR-2001; 2001US-0285325P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 12-MAY-2001; 2001US-0291134P.
PR 31-MAY-2001; 2001US-0294771P.
PR 08-JUN-2001; 2001US-0296965P.
PR 18-JUN-2001; 2001US-0299128P.
PR 12-JUL-2001; 2001US-0305063P.
PR 14-NOV-2001; 2001US-0332780P.
PR 04-JAN-2002; 2002US-0345221P.
PR 02-APR-2002; 2002US-00345221.
XX (CURA-) CURAGEN CORP.
PA Spyttek KA, Li L, Edinger SR, Ellerman K, Stone DJ, Malyankar UM;
XX Shinkets RA, Guo X, Anderson DW, Patturajan M, Berghs C, Gerlach V;
PI Taupier RJ, Pena CE, Padigaru M, Liu Y, Burgess CE, Miller CE;
PI Gusev VI, kekuda R, Gorman L, Zernusen BD, Baumgartner JC;
PI Tchernev VT, Vernet CAM, Smithson G, Heyes MP, Shenoy SG, Liu X;
PI Gangolli EA;

Db	1	CGTCAGTTCTTCTAGCAGGACCTGTCTCCCTTTTACTTCTTACCTCCACCTTTCCAGG	60
Qy	697	GCTTTCAAAGGAGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTCCACCAGAC	756
Db	61	GCTTTCAAAGGAGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTCCACCAGAC	120
Qy	757	TCAAAAGGACTCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATG	816
Db	121	TCAAAAGGACTCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATG	180
Qy	817	TCTGTCCCTTGGATAGGGTGAGGCTGAAGCACCCAGGGAGAAAATATGTCTTCTCTCGC	876
Db	181	TCTGTCCCTTGGATAGGGTGAGGCTGAAGCACCCAGGGAGAAAATATGTCTTCTCTCGC	240
Qy	877	CCTACCTCCTTTCCCATCCTAGACTGTCTTGAAGCCAGGGTCTGTAAACCTGACACTTTA	936
Db	241	CCTACCTCCTTTCCCATCCTAGACTGTCTTGAAGCCAGGGTCTGTAAACCTGACACTTTA	300
Qy	937	TATGTGTTCCACACATGTAAAGTAC	959
Db	301	TATGTGTTCCACACATGTAAAGNCC	323

Search completed: June 28, 2005, 19:33:46
 Job time : 918.543 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 18:22:52 ; Search time 295.707 Seconds
(without alignments)
14414.621 Million cell updates/sec

Title: US-08-731-499-4
Perfect score: 2605
Sequence: 1 CAAGCTCGAATAACCTC.....CTGGCGTTACCACTTAAT 2605

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2.6/prodata/1/ina/5A COMB.seq: *
2: /cgn2.6/prodata/1/ina/5B COMB.seq: *
3: /cgn2.6/prodata/1/ina/6A COMB.seq: *
4: /cgn2.6/prodata/1/ina/6B COMB.seq: *
5: /cgn2.6/prodata/1/ina/PTUS COMB.seq: *
6: /cgn2.6/prodata/1/ina/backfileseq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2605	100.0	2605	2	US-08-680-395-4
2	2605	100.0	2605	4	US-08-892-695-4
3	1076.2	41.3	1243	4	US-09-620-312D-930
4	476	18.3	495	4	US-09-621-976-308
5	453	17.4	487	4	US-09-621-976-307
6	444	17.0	494	4	US-09-621-976-1902
7	389.6	15.0	417	4	US-09-621-976-15507
8	386.4	14.8	400	4	US-09-621-976-15511
9	362	13.9	401	4	US-09-621-976-15506
10	328.2	12.6	383	4	US-09-621-976-15508
11	309	11.9	360	4	US-09-621-976-15509
12	269	10.3	269	4	US-09-702-705-1380
13	269	10.3	269	4	US-09-736-457-1380
14	269	10.3	269	4	US-09-614-124B-1380
15	269	10.3	269	4	US-09-671-325-1380
16	269	10.3	269	4	US-09-658-824-1380
17	227	8.7	241	4	US-09-471-276-611
18	209	8.0	261	4	US-09-621-976-15510
19	190	7.3	195	4	US-09-513-999C-433
20	156.4	6.0	1763	4	US-09-244-805-57
21	156.4	6.0	1824	4	US-09-244-805-5
22	145	5.6	5592	3	US-09-495-797-37
23	145	5.6	12479	4	US-09-318-138-13
24	145	5.6	12494	3	US-08-935-312-13
25	145	5.6	12494	3	US-08-848-760B-33
26	139.4	5.4	3219	4	US-09-509-800-3
27	138	5.3	4525	2	US-08-613-861-2

28	137.4	5.3	3385	4	US-09-509-800-1	Sequence 1, Appli
29	136	5.2	3699	3	US-08-646-538-6	Sequence 6, Appli
30	136	5.2	3699	3	US-09-503-222-6	Sequence 6, Appli
31	136	5.2	4965	3	US-08-675-566-22	Sequence 22, Appli
32	135	5.2	9299	3	US-09-097-319A-15	Sequence 15, Appli
33	135	5.2	9299	4	US-09-643-971-15	Sequence 15, Appli
34	134	5.1	6596	3	US-09-575-602-11	Sequence 11, Appli
35	134	5.1	6596	4	US-09-032-086-12	Sequence 12, Appli
36	133.8	5.1	3988	4	US-09-358-856C-12	Sequence 10881, A
37	133	5.1	902	4	US-09-270-767-10881	Sequence 1, Appli
38	133	5.1	6994	3	US-08-675-566-1	Sequence 1, Appli
39	133	5.1	7001	3	US-08-675-566-3	Sequence 3, Appli
40	133	5.1	9335	3	US-09-097-319A-19	Sequence 19, Appli
41	133	5.1	9335	4	US-09-643-971-19	Sequence 19, Appli
42	133	5.1	9408	3	US-09-097-319A-16	Sequence 16, Appli
43	133	5.1	9408	4	US-09-643-971-16	Sequence 16, Appli
44	130	5.0	4883	3	US-09-608-730B-21	Sequence 21, Appli
45	127.6	4.9	713	3	US-08-998-416-135	Sequence 135, App

ALIGNMENTS

RESULT 1
US-08-680-395-4
; Sequence 4, Application US/08680395
; Patent No. 5892010
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W.
; APPLICANT: Collins, Colin
; APPLICANT: Hwang, Soo-in
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowbel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: Genes from the 20q13 Amplicon and Their
; NUMBER OF INVENTION: Uses
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,395
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-0689000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2605 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..2605
; OTHER INFORMATION: /note= "cdna clone cc43 of 4 kb
; OTHER INFORMATION: transcript"
US-08-680-395-4

Query Match									
Best Local Similarity 100.0%; Score 2605; DB 2; Length 2605;									
Matches 2605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CAAGCTCGAAATTAACCCCTCACTAAAGGAAACAAAGCTGGAGCTCCACCGCGGTGGCGG	60						
Db	1	CAAGCTCGAAATTAACCCCTCACTAAAGGAAACAAAGCTGGAGCTCCACCGCGGTGGCGG	60						
Qy	61	CGGCTCTAGAACTAGTGGATCCCGCGGCTGAGGAAATTCGGCAGAGCTGGGCTACTAC	120						
Db	61	CGGCTCTAGAACTAGTGGATCCCGCGGCTGAGGAAATTCGGCAGAGCTGGGCTACTAC	120						
Qy	121	GATGCGGATGAGTTTCGAGTGGCGGCTGGCAGTATCGCTTCCACCCCTCTTTACGTTACA	180						
Db	121	GATGCGGATGAGTTTCGAGTGGCGGCTGGCAGTATCGCTTCCACCCCTCTTTACGTTACA	180						
Qy	181	ACCGAATGTGGACACTCGGCAAGCAGCTGGCCGCTGGTGTCTGGCTGCTGCTCTCT	240						
Db	181	ACCGAATGTGGACACTCGGCAAGCAGCTGGCCGCTGGTGTCTGGCTGCTGCTCTCT	240						
Qy	241	CTGCGGCTGCAAAACAGTTCAGCATGACGCTGATGGAGCTCAGGAGAGCCGCTCTT	300						
Db	241	CTGCGGCTGCAAAACAGTTCAGCATGACGCTGATGGAGCTCAGGAGAGCCGCTCTT	300						
Qy	301	CAACAACTCAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGATTTGTATTAGAGGA	360						
Db	301	CAACAACTCAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGATTTGTATTAGAGGA	360						
Qy	361	ACTGAGGAAGAAAGGAACTCTCGAGTGGTTGGATGAAGAGCAAGTCCAGCTTCTGATCAT	420						
Db	361	ACTGAGGAAGAAAGGAACTCTCGAGTGGTTGGATGAAGAGCAAGTCCAGCTTCTGATCAT	420						
Qy	421	GTGGCGGAGCCAGAAAGTGGGGAACTCATCTATCATGATGGGTTTCAGAGTGGGCA	480						
Db	421	GTGGCGGAGCCAGAAAGTGGGGAACTCATCTATCATGATGGGTTTCAGAGTGGGCA	480						
Qy	481	GAACAACTCCGCTTTACCTGTATGAACGTACTAATGGGGAAGACACAGAGGATGAGGA	540						
Db	481	GAACAACTCCGCTTTACCTGTATGAACGTACTAATGGGGAAGACACAGAGGATGAGGA	540						
Qy	541	GTTCACCGGCTGGATGAAGCCACTCTACTGGGGCTCTGAGGGCCCTACAGCAGGAGCA	600						
Db	541	GTTCACCGGCTGGATGAAGCCACTCTACTGGGGCTCTGAGGGCCCTACAGCAGGAGCA	600						
Qy	601	CAAGCCGAGATCATCTGTGAGGATGGCCGAGCGTCAAGTTCCTTAGCAGGGACC	660						
Db	601	CAAGCCGAGATCATCTGTGAGGATGGCCGAGCGTCAAGTTCCTTAGCAGGGACC	660						
Qy	661	TGTCTCCCTTTACTTCTTACCTCCACCTTTCCAGGGCTTTCAAAGGAGACAGACCCAG	720						
Db	661	TGTCTCCCTTTACTTCTTACCTCCACCTTTCCAGGGCTTTCAAAGGAGACAGACCCAG	720						
Qy	721	TGTCCCCCAAAGACTGGATCTGTGACTCCACAGACTCAAAGGACTCCAGTCTGAAAG	780						
Db	721	TGTCCCCCAAAGACTGGATCTGTGACTCCACAGACTCAAAGGACTCCAGTCTGAAAG	780						
Qy	781	CTGGGACCTGGGGATGGGTTTCTCAACCCATATGTCTGCTCCCTTGGATAGGGTGAGGC	840						
Db	781	CTGGGACCTGGGGATGGGTTTCTCAACCCATATGTCTGCTCCCTTGGATAGGGTGAGGC	840						
Qy	841	TGAAGCACAGGAGGAAATAATGTCTTCTCTGCGCCCTACCTCTCTTCCCATCTAGAC	900						
Db	841	TGAAGCACAGGAGGAAATAATGTCTTCTCTGCGCCCTACCTCTCTTCCCATCTAGAC	900						
Qy	901	TGTCTCTGAGCCAGGGTCTGTAACCTGACACTTTATATGTGTTTACACATGTAAGTACA	960						
Db	901	TGTCTCTGAGCCAGGGTCTGTAACCTGACACTTTATATGTGTTTACACATGTAAGTACA	960						
Qy	961	TACACATGCGGCTGACGACATGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020						
Db	961	TACACATGCGGCTGACGACATGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020						

Qy	1021	TGTTGCTCCCTTCTCAGCTGGTCTGGATCCCTTCTTAGGGGATGGGGAAAGCCCTGGC	1080
Db	1021	TGTTGCTCCCTTCTCAGCTGGTCTGGATCCCTTCTTAGGGGATGGGGAAAGCCCTGGC	1080
Qy	1081	TGACGAGCCCTTCCAGGCAATATGAAATAGGAGGCCACCGGGCTGGCAGTGAGAGT	1140
Db	1081	TGACGAGCCCTTCCAGGCAATATGAAATAGGAGGCCACCGGGCTGGCAGTGAGAGT	1140
Qy	1141	GTGGCCACACACCGGATTTATGATATAAATCTCACTCCCAAAAAAAGGAAAAA	1200
Db	1141	GTGGCCACACACCGGATTTATGATATAAATCTCACTCCCAAAAAAAGGAAAAA	1200
Qy	1201	CTGAGACTAGTCTCTCTCTCGAGAACTAGTCTCGAGTTTTTTTTTTTTTTTTTTTT	1260
Db	1201	CTGAGACTAGTCTCTCTCTCGAGAACTAGTCTCGAGTTTTTTTTTTTTTTTTTTTT	1260
Qy	1261	TTTTTTTTTTTTTTTTTTTTTTGCTTAAAGATTTATTTATTTGTTTCTCTTTACAGTGTCC	1320
Db	1261	TTTTTTTTTTTTTTTTTTTTTTGCTTAAAGATTTATTTATTTGTTTCTCTTTACAGTGTCC	1320
Qy	1321	ACTTTTCTTACTTAACTACTTTTCCAGTCTCAGAGCCAGAGGAAAAAAGAGAC	1380
Db	1321	ACTTTTCTTACTTAACTACTTTTCCAGTCTCAGAGCCAGAGGAAAAAAGAGAC	1380
Qy	1381	CATGAATCTTCTCTCCAGATTTAAAGTACACATTTTGGAAAAAGATTTGGAAAACTTT	1440
Db	1381	CATGAATCTTCTCTCCAGATTTAAAGTACACATTTTGGAAAAAGATTTGGAAAACTTT	1440
Qy	1441	CTGAAAAAGTTGACTGAAACTCCAAAACCATGCCATATTTGTTGATGTTGCTCATGAA	1500
Db	1441	CTGAAAAAGTTGACTGAAACTCCAAAACCATGCCATATTTGTTGATGTTGCTCATGAA	1500
Qy	1501	AATTTGTTAAAGCTGTTCTAGATAAAGAGAGTCTCAAGTTTTTTGTACAGCTACACAT	1560
Db	1501	AATTTGTTAAAGCTGTTCTAGATAAAGAGAGTCTCAAGTTTTTTGTACAGCTACACAT	1560
Qy	1561	AGTACAAGGCTCCCTATGATGATTTCTTCTAGGACGAAATAATGTAATTTTTTTCAGTT	1620
Db	1561	AGTACAAGGCTCCCTATGATGATTTCTTCTAGGACGAAATAATGTAATTTTTTTCAGTT	1620
Qy	1621	TCTGTTTTTAACTCTCTCGATCTCAGAGTTGACTGATTTAAACAACCTTCTCATGCAACA	1680
Db	1621	TCTGTTTTTAACTCTCTCGATCTCAGAGTTGACTGATTTAAACAACCTTCTCATGCAACA	1680
Qy	1681	GAGATAAGGCACTCATATTTTATAAATATATGACCAAACTATTTTGGAAAACTTAT	1740
Db	1681	GAGATAAGGCACTCATATTTTATAAATATATGACCAAACTATTTTGGAAAACTTAT	1740
Qy	1741	CTATTGGAGACAAATATGCTGGAATAAGCAATAATTTATTTTCTCAATGTCTGTGC	1800
Db	1741	CTATTGGAGACAAATATGCTGGAATAAGCAATAATTTATTTTCTCAATGTCTGTGC	1800
Qy	1801	TAACTCAATGACTTAGAATGCTTTGCTATATTTTGGCTCTATGCTCAACACACTGGC	1860
Db	1801	TAACTCAATGACTTAGAATGCTTTGCTATATTTTGGCTCTATGCTCAACACACTGGC	1860
Qy	1861	TTTTTTTTTAGCTTTGAAACAGCCAACTGCTTCTGCTCAGGACAGATATTTTGGGA	1920
Db	1861	TTTTTTTTTAGCTTTGAAACAGCCAACTGCTTCTGCTCAGGACAGATATTTTGGGA	1920
Qy	1921	CTTCTCTTAAAGAAATCTATTTTCTTAAATCTTTATCTGGGTAACTTAGTTTTTATCAACA	1980
Db	1921	CTTCTCTTAAAGAAATCTATTTTCTTAAATCTTTATCTGGGTAACTTAGTTTTTATCAACA	1980
Qy	1981	CTTCAGATCTCGCCGTAAAGAACTTTCTTATAGAGCCGTGTGACACTGTCTCTTTC	2040
Db	1981	CTTCAGATCTCGCCGTAAAGAACTTTCTTATAGAGCCGTGTGACACTGTCTCTTTC	2040
Qy	2041	TCCACATACCTACACAGCACACATCTAGACTAGATTAGAACCTCTCTGTTTTTTTTTTCA	2100
Db	2041	TCCACATACCTACACAGCACACATCTAGACTAGATTAGAACCTCTCTGTTTTTTTTTTCA	2100
Qy	2101	TACTTTTCTCTATCATGCTTCCCTCCATTAATAATTTTTTATTATGTGTGTAATGTCTG	2160

Db 2101 TACTTTCTATCATGCTTCCCTCCATTATATATATTTTATATGTTGTAATGTCGTG 2160
QY 2161 CCCAAGTCAGTTTCTCTCACTAAATATATAAATCCCGTAAAGCTGGGATCCTTCCCAATTTT 2220
Db 2161 CCCAAGTCAGTTTCTCTCACTAAATATATAAATCCCGTAAAGCTGGGATCCTTCCCAATTTT 2220
QY 2221 GATCACCATTAGTACAGTAGGAACACAGTAAAGATTCAATTGGTATTTTGTGGAATGAAT 2280
Db 2221 GATCACCATTAGTACAGTAGGAACACAGTAAAGATTCAATTGGTATTTTGTGGAATGAAT 2280
QY 2281 GAATGAATTTTGTCTAGTAAAGCTTGGGGAAACCCAGGTGAGAGAGCCCTAGAAAGCA 2340
Db 2281 GAATGAATTTTGTCTAGTAAAGCTTGGGGAAACCCAGGTGAGAGAGCCCTAGAAAGCA 2340
QY 2341 GGTTCGAATCCAAAGGCTAGATAGACTTAGTGTGTACTCAAGAAAGGCTAGCCTGAAATATAA 2400
Db 2341 GGTTCGAATCCAAAGGCTAGATAGACTTAGTGTGTACTCAAGAAAGGCTAGCCTGAAATATAA 2400
QY 2401 GGTTCGAATTTAGTACAGAAATAGTCAAGATAGTCAAGATAGTGGCAAGACAGAGTCTCTGTGTC 2460
Db 2401 GGTTCGAATTTAGTACAGAAATAGTCAAGATAGTCAAGATAGTGGCAAGACAGAGTCTCTGTGTC 2460
QY 2461 GAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGGGGGCCCGGTACCCCAATTC 2520
Db 2461 GAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGGGGGCCCGGTACCCCAATTC 2520
QY 2521 GCCCTATAGTGTGCTATTACAAATTAATCACTGGCCGCTGCTTTTACAAAGCTCGTGTGGA 2580
Db 2521 GCCCTATAGTGTGCTATTACAAATTAATCACTGGCCGCTGCTTTTACAAAGCTCGTGTGGA 2580
QY 2581 AAACCTGGCTTACCCCACTTAAT 2605
Db 2581 AAACCTGGCTTACCCCACTTAAT 2605

RESULT 2

US-08-892-695-4
; Sequence 4, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2605
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cc43
US-08-892-695-4

Query Match 100.0%; Score 2605; DB 4; Length 2605;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCTCGAAATTAACCCCTCACTAAAGGAAACAAAGAGCTGGAGCTCCACCGCGTGGCGG 60
|||||

Db 1 CAAGCTCGAAATTAACCCCTCACTAAAGGAAACAAAGAGCTGGAGCTCCACCGCGGTGGCGG 60
QY 61 CCGCTCTAGAACTAGTGGATCCCGGGCTGAGGAATTCGCGACAGAGCTGGGCTACTAC 120
Db 61 CCGCTCTAGAACTAGTGGATCCCGGGCTGAGGAATTCGCGACAGAGCTGGGCTACTAC 120
QY 121 GATGGGATGAGTTTCGAGTGGCCGCTGGCAGATATGCTTCCACCCCTTCTTACGTTACA 180
Db 121 GATGGGATGAGTTTCGAGTGGCCGCTGGCAGATATGCTTCCACCCCTTCTTACGTTACA 180
QY 181 ACCGAATGTGGACACTCGSCAGAAGCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTT 240
Db 181 ACCGAATGTGGACACTCGSCAGAAGCAGCTGCGCTGCTGCTGCTGCTGCTGCTT 240
QY 241 CTGCGCCCTGCAACAAAGTCCAGCATGACGCTGATGGAAGCTCAGAGAGAGCCCGCTCTT 300
Db 241 CTGCGCCCTGCAACAAAGTCCAGCATGACGCTGATGGAAGCTCAGAGAGAGCCCGCTCTT 300
QY 301 CAAACAGTCAAGCTACAGCGAAAGCTTCTGTGAGTGCATCCAGATTTGTTATAGAGA 360
Db 301 CAAACAGTCAAGCTACAGCGAAAGCTTCTGTGAGTGCATCCAGATTTGTTATAGAGA 360
QY 361 ACTGAGGAAGAAAGGAAACCTCGAGTGGTGGATAGAGCAAGTCCAGCTTCTGTATCAT 420
Db 361 ACTGAGGAAGAAAGGAAACCTCGAGTGGTGGATAGAGCAAGTCCAGCTTCTGTATCAT 420
QY 421 GTGGCGGAGCCAGAAAGTGGGGAAACTCATCTATCAGTGGGTTTCCAGAGAGTGCCCA 480
Db 421 GTGGCGGAGCCAGAAAGTGGGGAAACTCATCTATCAGTGGGTTTCCAGAGAGTGCCCA 480
QY 481 GAAACCTCCGCTTTTACCTTATGAACTGAACTAAATGGGGAAGACACAGAGGATGAGGA 540
Db 481 GAAACCTCCGCTTTTACCTTATGAACTGAACTAAATGGGGAAGACACAGAGGATGAGGA 540
QY 541 GTTCCACGCGCTGGATGAAGCCACTCTACTGCGGGCTCTGCGAGGCTTACAGAGAGGCA 600
Db 541 GTTCCACGCGCTGGATGAAGCCACTCTACTGCGGGCTCTGCGAGGCTTACAGAGAGGCA 600
QY 601 CAAGCCGAGATCATCACTGTGAGGATGGCGAGGCGTCAAGTCTTCTAGCAGGAGCC 660
Db 601 CAAGCCGAGATCATCACTGTGAGGATGGCGAGGCGTCAAGTCTTCTAGCAGGAGCC 660
QY 661 TGTCTCCCTTTACTTCTTACCTTCCACCTTCCAGGGCTTTCAAAGGAGACAGACCCAG 720
Db 661 TGTCTCCCTTTACTTCTTACCTTCCACCTTCCAGGGCTTTCAAAGGAGACAGACCCAG 720
QY 721 TGTCCCCCAAGAGCTGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCCCTGAAGG 780
Db 721 TGTCCCCCAAGAGCTGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCCCTGAAGG 780
QY 781 CTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGCTCCCTTGGATAGGGTGAGGC 840
Db 781 CTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGCTCCCTTGGATAGGGTGAGGC 840
QY 841 TGAAGCACACAGGAGAAATATGTGCTTCTGCGCTTCTGCTTCCCTTCCCATCTTAGAC 900
Db 841 TGAAGCACACAGGAGAAATATGTGCTTCTGCGCTTCTGCTTCCCTTCCCATCTTAGAC 900
QY 901 TGTCTCTGAGCCAGGCTGTAAACCTGACATTTATATGTGCTTCCACATCTAAGTACA 960
Db 901 TGTCTCTGAGCCAGGCTGTAAACCTGACATTTATATGTGCTTCCACATCTAAGTACA 960
QY 961 TACACATATGGCCCTGACAGCAGCATGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Db 961 TACACATATGGCCCTGACAGCAGCATGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
QY 1021 TGTTCCTTCCCTTCTCAGGCTGGTCTGGATCCTTCTTAGGGGATGGGGAGAGCCCTGGC 1080
Db 1021 TGTTCCTTCCCTTCTCAGGCTGGTCTGGATCCTTCTTAGGGGATGGGGAGAGCCCTGGC 1080
QY 1081 TGCAGGAGCTTCCAGGCAATATGAAGTAGAGAGGCCCCACCGGCTGGCAGTGAGGT 1140
Db 1081 TGCAGGAGCTTCCAGGCAATATGAAGTAGAGAGGCCCCACCGGCTGGCAGTGAGGT 1140

1141 GTGGCCCAACCGATTATATGATATATAAACTCAACTCCCAAAAAAAAAAAAAAAAAA 1200
Db
1141 GTGGCCCAACCGATTATATGATATATAAACTCAACTCCCAAAAAAAAAAAAAAAAAA 1200
1201 CTGAGACTAGTTCTCTCTCTCGAGAACTAGTCTCGAGTCTTTTCTTTTCTTTTCTTTT 1260
Db
1201 CTGAGACTAGTTCTCTCTCTCGAGAACTAGTCTCGAGTCTTTTCTTTTCTTTTCTTTT 1260
1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1320
Db
1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1320
1321 ACTTTTCTCTAACTAACTACTTTTCCAGTCTCAGAACCCAGAGGGGAAAAAAAAAGAC 1380
Db
1321 ACTTTTCTCTAACTAACTACTTTTCCAGTCTCAGAACCCAGAGGGGAAAAAAAAAGAC 1380
1381 CATGAATCTTCTCTCCAGATTAAGTACACACTTTTGGAAAAACAGATTGGAAAACTTTT 1440
Db
1381 CATGAATCTTCTCTCCAGATTAAGTACACACTTTTGGAAAAACAGATTGGAAAACTTTT 1440
1441 CTGAAAAAGTTGACTGAAACTCCAAACCAATGCCATATTTGTTGATGTTGCTCATGAA 1500
Db
1441 CTGAAAAAGTTGACTGAAACTCCAAACCAATGCCATATTTGTTGATGTTGCTCATGAA 1500
1501 AATTGTTAAAAACCTGTTCTAGATAAAGAAAGCTCTCAAGTTTGTACAGCCTACACAT 1560
Db
1501 AATTGTTAAAAACCTGTTCTAGATAAAGAAAGCTCTCAAGTTTGTACAGCCTACACAT 1560
1561 AGTACAAGGGTCCCTATGATGATCTTCTGTAGGACGAAATAATGTAATTTTTTCAGTT 1620
Db
1561 AGTACAAGGGTCCCTATGATGATCTTCTGTAGGACGAAATAATGTAATTTTTTCAGTT 1620
1621 TCTGTTTATAACTCTCTCGATCTCAGAGTTGACTGATTAAGAACCTACTCATGCAACA 1680
Db
1621 TCTGTTTATAACTCTCTCGATCTCAGAGTTGACTGATTAAGAACCTACTCATGCAACA 1680
1681 GAGATAAAGCACTCATATTTTATAAATATATGAGCAAACTATTTTGGAAATCTTAT 1740
Db
1681 GAGATAAAGCACTCATATTTTATAAATATATGAGCAAACTATTTTGGAAATCTTAT 1740
1741 CTATTGGAGACACAATATGCTGGACTAAAGCAATAATTTATTTATCTCAATGCTGTGC 1800
Db
1741 CTATTGGAGACACAATATGCTGGACTAAAGCAATAATTTATTTATCTCAATGCTGTGC 1800
1801 TAACTCAATGACTTTAGAAATGCTTTGCTATATTTTGGCCTCTATGCTCAACCACTGGC 1860
Db
1801 TAACTCAATGACTTTAGAAATGCTTTGCTATATTTTGGCCTCTATGCTCAACCACTGGC 1860
1861 TTTCTTTTGTAGCTTTGAAACAGCCAAACTGCTTCTGCTCAGGACCAAGATATTTGGGA 1920
Db
1861 TTTCTTTTGTAGCTTTGAAACAGCCAAACTGCTTCTGCTCAGGACCAAGATATTTGGGA 1920
1921 CTTCTCTTAAAGAAATCTATTTCTTTAACTTTCTTTATCTGGGTAACCTAGTTTATCAACA 1980
Db
1921 CTTCTCTTAAAGAAATCTATTTCTTTAACTTTCTTTATCTGGGTAACCTAGTTTATCAACA 1980
1981 CTTGAGATCTGCGGTAAAACTCTTTATAGAAAGCCTGTATGACACTGCTCTCTTC 2040
Db
1981 CTTGAGATCTGCGGTAAAACTCTTTATAGAAAGCCTGTATGACACTGCTCTCTTC 2040
2041 TCCAACTACTCACCAGACACATGCTAGACTAGATTAGAACTCTGTTTCTTTTCTTTTCA 2100
Db
2041 TCCAACTACTCACCAGACACATGCTAGACTAGATTAGAACTCTGTTTCTTTTCTTTTCA 2100
2101 TACTTTTCTATCATGCTTCCCTCCATTAATAATTTTTTATGTTGTAATGCTG 2160
Db
2101 TACTTTTCTATCATGCTTCCCTCCATTAATAATTTTTTATGTTGTAATGCTG 2160
2161 CCCAAGTCAGTTTCTCCTCACTAAACTATAAAGTAAAGTGGATCCTTCCAAATTTT 2220
Db
2161 CCCAAGTCAGTTTCTCCTCACTAAACTATAAAGTAAAGTGGATCCTTCCAAATTTT 2220

2221 GATCACCACCTTAGTACAGTAGGAAACACAGTAAAGATTCAATTGGTATTTGTGAATGAAT 2280
Db
2221 GATCACCACCTTAGTACAGTAGGAAACACAGTAAAGATTCAATTGGTATTTGTGAATGAAT 2280
2281 GAATGAATTTGTTTCTAGTAAAGTCTGGGGGAAACCCAGGTGAGAAAGACCTTAGAAAGCA 2340
Db
2281 GAATGAATTTGTTTCTAGTAAAGTCTGGGGGAAACCCAGGTGAGAAAGACCTTAGAAAGCA 2340
2341 GGTGGAATCCAAAGCTAGATAGACTTAGTGTCTCAAGAAAGGGTAGCCTGAAATATAA 2400
Db
2341 GGTGGAATCCAAAGCTAGATAGACTTAGTGTCTCAAGAAAGGGTAGCCTGAAATATAA 2400
2401 GGTTCAAATATTAGTCAAGAAATAGTCAAGACATGGGCAAGACAGAGTGTCTGCTGCC 2460
Db
2401 GGTTCAAATATTAGTCAAGAAATAGTCAAGACATGGGCAAGACAGAGTGTCTGCTGCC 2460
2461 GAATTCGATATCAAGCTTTATCGATACCTCGAGGGGGGGCCGGTACCCAAATTC 2520
Db
2461 GAATTCGATATCAAGCTTTATCGATACCTCGAGGGGGGGCCGGTACCCAAATTC 2520
2521 GCCCTATAGTCAAGCTGATTTACAATTTCACTGGCGTCTTTTACAACTGCTGACTGGGA 2580
Db
2521 GCCCTATAGTCAAGCTGATTTACAATTTCACTGGCGTCTTTTACAACTGCTGACTGGGA 2580
2581 AAACCTGGCGTTACCCAACTTAAT 2605
Db
2581 AAACCTGGCGTTACCCAACTTAAT 2605

RESULT 3
US-09-620-312D-930
; Sequence 930, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aildong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 930
; LENGTH: 1243
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (172)..(702)
US-09-620-312D-930

Query Match 41.3%; Score 1076.2; DB 4; Length 1243;
Best Local Similarity 99.7%; Pred. No. 4e-247;

Matches 1078; Conservative 0; Mismatches 3; Indels 0; Gaps 0;					
QY	113	GCTACTACGATGGCGATGATTTTCGATGGCCGTCGATGATCGTTCCACCCCTTTT	172		
DB	163	CGTTCGACGATGGCGATGATTTTCGATGGCCGTCGATGATCGTTCCACCCCTTTT	222		
QY	173	AGTTTACAACCGAATGTGACACACTCGGCGAGAGCAGCTGGCCGCTTGTGCTCGTGTTC	232		
DB	223	ACGTTACAACCGAATGTGACACACTCGGCGAGAGCAGCTGGCCGCTTGTGCTCGTGTTC	282		
QY	233	CTGTCTCTCTCGCGCTCGACAAACAGTCCAGCATGACGGTGTGATGAAGCTCAGGAGAC	292		
DB	283	CTGTCTCTCTCGCGCTCGACAAACAGTCCAGCATGACGGTGTGATGAAGCTCAGGAGAC	342		
QY	293	CGCTCTTTCAACACGTCGAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCAGATTGTA	352		
DB	343	CGCTCTTTCAACACGTCGAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCAGATTGTA	402		
QY	353	TTAGAGGAAGTGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAGCTTC	412		
DB	403	TTAGAGGAAGTGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAGCTTC	462		
QY	413	CTGATCATGTGGCGAGGCGAGAGAAATGGGGGAACTCATCTATCAGTGGGTTCACAGG	472		
DB	463	CTGATCATGTGGCGAGGCGAGAGAAATGGGGGAACTCATCTATCAGTGGGTTCACAGG	522		
QY	473	AGTGGCCAGAACAACTCCGCTTTTACCTCTGTATGAATGACTAAATGGGGAACACAGAG	532		
DB	523	AGTGGCCAGAACAACTCCGCTTTTACCTCTGTATGAATGACTAAATGGGGAACACAGAG	582		
QY	533	GATGAGGATTCACGGGCTGATGAGCACTCTATCGCGGCTCTGAGGCCCTACAG	592		
DB	583	GATGAGGATTCACGGGCTGATGAGCACTCTATCGCGGCTCTGAGGCCCTACAG	642		
QY	593	CAGGAGCAACAGCCGAGATCATCTGTCAGGATGGCCGAGGCTCAAGTTCCTCTAG	652		
DB	643	CAGGAGCAACAGCCGAGATCATCTGTCAGGATGGCCGAGGCTCAAGTTCCTCTAG	702		
QY	653	CAGGAGCACTCTCCCTTTACTTTTACTTCTTACCTCCACCTTTCCAGGGCTTTCAAAAGGAGAC	712		
DB	703	CAGGAGCACTCTCCCTTTACTTTTACTTCTTACCTCCACCTTTCCAGGGCTTTCAAAAGGAGAC	762		
QY	713	AGACCCAGTGTCCCCAAAGACTGAGTCTGTGATCTCCACGACTCAAAAGACTCCAGT	772		
DB	763	AGACCCAGTGTCCCCAAAGACTGAGTCTGTGATCTCCACGACTCAAAAGACTCCAGT	822		
QY	773	CCTGAAGGCTGGGACCTGGGATGGTCTTCTCACCCCATATGCTGCTCCCTCGATAG	832		
DB	823	CCTGAAGGCTGGGACCTGGGATGGTCTTCTCACCCCATATGCTGCTCCCTCGATAG	882		
QY	833	GGTGAGGCTGAAGCACACGAGGAGAAATATGTGCTTCTTCGCCCTACCTCTTTCCCA	892		
DB	883	GGTGAGGCTGAAGCACACGAGGAGAAATATGTGCTTCTTCGCCCTACCTCTTTCCCA	942		
QY	893	TCCTAGACTGTCTTTGAGCCAGGGCTGTGTAACCTGACACTTTATATGTGTTTCAACATG	952		
DB	943	TCCTAGACTGTCTTTGAGCCAGGGCTGTGTAACCTGACACTTTATATGTGTTTCAACATG	1002		
QY	953	TAAGTACATACACATGCGGCTGACGACATGCTTCTGCTCTCTCCCTCCGACCCCT	1012		
DB	1003	TAAGTACATACACATGCGGCTGACGACATGCTTCTGCTCTCTCCCTCCGACCCCT	1062		
QY	1013	TTAGTGTCTGTGCTCTCTCTCAGGCTGGTGTGATCTTCTCTAGGGGATGGGGAA	1072		
DB	1063	TTAGTGTCTGTGCTCTCTCTCAGGCTGGTGTGATCTTCTCTAGGGGATGGGGAA	1122		
QY	1073	GCCCTGGCTGACGAGGCTTCCAGGCAATATGAAGATAGAGGCCCAAGGCCCTGGCAG	1132		
DB	1123	GCCCTGGCTGACGAGGCTTCCAGGCAATATGAAGATAGAGGCCCAAGGCCCTGGCAG	1182		
QY	1133	TCAGAGGCTGGCCCCACACCGATTTATGATATTTAAATCTCACTCCCAAAAAA	1192		
DB	1183	TCAGAGGCTGGCCCCACACCGATTTATGATATTTAAATCTCACTCCCAAAAAA	1242		

QY 1193 A 1193
DB 1243 A 1243

RESULT 4
US-09-621-976-308
; Sequence 308, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 308
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..495
; NAME/KEY: sig peptide
; LOCATION: 28..156
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4
; OTHER INFORMATION: seq AAWCSLVLSFCRL/HK
US-09-621-976-308

Query Match 18.3%; Score 476; DB 4; Length 495;
Best Local Similarity 99.6%; Pred. No. 7.3e-104; Indels 0; Gaps 0;
Matches 476; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	109	CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTCGAGTATCGCTTCCACCCCTT	168
DB	15	CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTCGAGTATCGCTTCCACCCCTT	74
QY	169	CTTTACGTTTACAACCGAATGTGGACACTCGGCGAAGCAGCTGGCCGCTGTGTCTCGT	228
DB	75	CTTTACGTTTACAACCGAATGTGGACACTCGGCGAAGCAGCTGGCCGCTGTGTCTCGT	134
QY	229	GGTCTGTCTTCTGCGCGCTGCACAAACAGTCCAGCATGACGCTGATGGAAGCTCAGGA	288
DB	135	GGTCTGTCTTCTGCGCGCTGCACAAACAGTCCAGCATGACGCTGATGGAAGCTCAGGA	194
QY	289	GAGCCCGCTCTTCAACACGCTACAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGAT	348
DB	195	GAGCCCGCTCTTCAACACGCTACAGCTACAGCGAAAGSTTCTGTGGAGTCGATCCAGAT	254
QY	349	TGTATTAGAGGAAGTGAAGAAAGGAAACCTCGAGTGGTGGATAAGAGCAAGTCCAG	408
DB	255	TGTATTAGAGGAAGTGAAGAAAGGAAACCTCGAGTGGTGGATAAGAGCAAGTCCAG	314
QY	409	CTTCTGTATCATGTGGCCGAGCCAGAAATGGGGGAACTCATCTATCAGTGGGTTC	468
DB	315	CTTCTGTATCATGTGGCCGAGCCAGAAATGGGGGAACTCATCTATCAGTGGGTTC	374
QY	469	CAGGATGGCCAGAACAACTCCGCTCTTTACCTGTATGAACCTGACTAATGGGGAAGACAC	528
DB	375	CAGGATGGCCAGAACAACTCCGCTCTTTACCTGTATGAACCTGACTAATGGGGAAGACAC	434
QY	529	AGAGGATCAGGAGTTCACAGGCTGGATGAAGCCACTCTATCTGGGGCTCTGCAGGCC	586
DB	435	AGAGGATCAGGAGTTCACAGGCTGGATGAAGCCACTCTATCTGGGGCTCTGCAGGCC	492

RESULT 5
US-09-621-976-307


```
FEATURE:
NAME/KEY: misc_feature
LOCATION: 261,327,360
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-15507

Query Match
Best Local Similarity 15.0%; Score 389.6; DB 4; Length 417;
Matches 411; Conservative 2; Mismatches 4; Indels 2; Gaps 2;

QY 113 GCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCACCCCTCTCTTT 172
DB 1 GCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCACCCCTCTCTTT 60

QY 173 ACGTTACAAACCGAATGTGACACTCGGAGAGACAGTGGCCGCTGGTGTGCTGCTGGTC 232
DB 61 ACGTTACAAACCGAATGTGACACTCGGAGAGACAGTGGCCGCTGGTGTGCTGCTGGTC 120

QY 233 CTGTCTCTTCGCGCTCGACAAACAGTCCAGCATGACGGTGTATGGAAGCTCAGGAGAGC 292
DB 121 CTGTCTCTTCGCGCTCGACAAACAGTCCAGCATGACGGTGTATGGAAGCTCAGGAGAGC 180

QY 293 CGGCTCTTCAACAAACGTCAGTCTGAGCGAAAGCTTCCTGTGGAGTCGATCCAGATTGTA 352
DB 181 CGGCTCTTCAACAAACGTCAGTCTGAGCGAAAGCTTCCTGTGGAGTCGATCCAGATTGTA 239

QY 353 TTAGAGGAACCTGAGGAAGAAAGGAAACCTCGAGTGGTTCGATAAGAGCAAGTCCAGCTTC 412
DB 240 TTAGAGGAACCTGAGGAAGAAAGGAAACCTCGAGTGGTTCGATAAGAGCAAGTCCAGCTTC 299

QY 413 CTGATCATGTGCGGAGGCGCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTCAGG 472
DB 300 CTGATCATGTGCGGAGGCGCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTCAGG 358

QY 473 AGTGGCCAGAACAACTCCGCTCTTTACCTGTATGAAGTCACTAAATGGGGAAGACACAGA 531
DB 359 ANTGGCCAGAACAACTCCGCTCTTTACCTGTATGAAGTCACTAAATGGGGAAGACACAGA 417

RESULT 8
US-09-621-976-15511
Sequence 15511, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 15511
LENGTH: 400
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 284
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-15511

Query Match
Best Local Similarity 14.8%; Score 386.4; DB 4; Length 400;
Matches 398; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 116 ACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCACCCCTCTTTTACG 175
DB 1 ACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCACCCCTCTTTTACG 60

QY 176 TTACAAACCGAATGTGACACTCGGAGAGACAGTGGCCGCTGGTGTGCTGCTGGTCCTG 235
DB 176 TTACAAACCGAATGTGACACTCGGAGAGACAGTGGCCGCTGGTGTGCTGCTGGTCCTG 235

FEATURE:
NAME/KEY: misc_feature
LOCATION: 261,327,360
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-15507

Query Match
Best Local Similarity 15.0%; Score 389.6; DB 4; Length 417;
Matches 411; Conservative 2; Mismatches 4; Indels 2; Gaps 2;

QY 113 GCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCACCCCTCTCTTT 172
DB 1 GCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCACCCCTCTCTTT 60

QY 173 ACGTTACAAACCGAATGTGACACTCGGAGAGACAGTGGCCGCTGGTGTGCTGCTGGTC 232
DB 61 ACGTTACAAACCGAATGTGACACTCGGAGAGACAGTGGCCGCTGGTGTGCTGCTGGTC 120

QY 233 CTGTCTCTTCGCGCTCGACAAACAGTCCAGCATGACGGTGTATGGAAGCTCAGGAGAGC 292
DB 121 CTGTCTCTTCGCGCTCGACAAACAGTCCAGCATGACGGTGTATGGAAGCTCAGGAGAGC 180

QY 293 CGGCTCTTCAACAAACGTCAGTCTGAGCGAAAGCTTCCTGTGGAGTCGATCCAGATTGTA 352
DB 181 CGGCTCTTCAACAAACGTCAGTCTGAGCGAAAGCTTCCTGTGGAGTCGATCCAGATTGTA 239

QY 353 TTAGAGGAACCTGAGGAAGAAAGGAAACCTCGAGTGGTTCGATAAGAGCAAGTCCAGCTTC 412
DB 240 TTAGAGGAACCTGAGGAAGAAAGGAAACCTCGAGTGGTTCGATAAGAGCAAGTCCAGCTTC 299

QY 413 CTGATCATGTGCGGAGGCGCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTCAGG 472
DB 300 CTGATCATGTGCGGAGGCGCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTCAGG 358

QY 473 AGTGGCCAGAACAACTCCGCTCTTTACCTGTATGAAGTCACTAAATGGGGAAGACACAGA 531
DB 359 ANTGGCCAGAACAACTCCGCTCTTTACCTGTATGAAGTCACTAAATGGGGAAGACACAGA 417

RESULT 9
US-09-621-976-15506
Sequence 15506, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 15506
LENGTH: 401
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 86,316,359,366
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-15506

Query Match
Best Local Similarity 13.9%; Score 362; DB 4; Length 401;
Matches 392; Conservative 5; Mismatches 4; Indels 3; Gaps 3;

QY 113 GCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCACCCCTCTCTTT 172
DB 1 GCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCACCCCTCTCTTT 60

QY 173 ACGTTACAAACCGAATGTGACACTCGGAGAGACAGTGGCCGCTGGTGTGCTGCTGGTC 232
DB 61 ACGTTACAAACCGAATGTGACACTSNGCAGAGACAGTGGCCGCTGGTGTGCTGCTGGTC 120

QY 233 CTGTCTCTTCGCGCTCGACAAACAGTCCAGCATGACGGTGTATGGAAGCTCAGGAGAGC 292
DB 121 CTGTCTCTTCGCGCTCGACAAACAGTCCAGCATGACGGTGTATGGAAGCTCAGGAGAGC 180

QY 293 CGGCTCTTCAACAAACGTCAGTCTGAGCGAAAGCTTCCTGTGGAGTCGATCCAGATTGTA 352
DB 181 CGGCTCTTCAACAAACGTCAGTCTGAGCGAAAGCTTCCTGTGGAGTCGATCCAGATTGTA 239

QY 353 TTAGAGGAACCTGAGGAAGAAAGGAAACCTCGAGTGGTTCGATAAGAGCAAGTCCAGCTTC 412
DB 240 TTAGAGGAACCTGAGGAAGAAAGGAAACCTCGAGTGGTTCGATAAGAGCAAGTCCAGCTTC 299

QY 413 CTGATCATGTGCGGAGGCGCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTCAGG 472
DB 359 ANTGGCCAGAACAACTCCGCTCTTTACCTGTATGAAGTCACTAAATGGGGAAGACACAGA 417
```

```
Db 300 CTGATCATGTGGCGGANGCCA-AAGAAATGGGGGAAACTCATCTACGTGGTTTCCA-G 357
Qy 473 AGTGGCCGAACAACCTCCGTCTTTTACCCCTGTATGAACGACTAA 516
Db 358 ANTGGCCANAACAACCTCCGTCTTTTACCCCTGTATGAACGACTAA 401

RESULT 10
US-09-621-976-15508
; Sequence 15508, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15508
; LENGTH: 383
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 183..184
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-15508

Query Match 12.6%; Score 328.2; DB 4; Length 383;
Best Local Similarity 96.8%; Pred. No. 1.4e-68;
Matches 359; Conservative 6; Mismatches 3; Indels 3; Gaps 3;

Qy 109 CTGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCCAACCCCTT 168
Db 15 CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCCAACCCCTT 74

Qy 169 CTTTACGTT-ACAACCGAATGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGC 227
Db 75 CTTTAMGKTAACAACGGAATGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGC 134

Qy 228 TGGTCTGTCTCTTCGCGCCCTGCACAAACAGTCAGCATGACGCTGATGGAAGCTCAGG 287
Db 135 TGGTCTGTCTCTTCGCGCCCTGCACAAACAGTCAGCATGACGCTVHNGAGCTCAGG 194

Qy 288 AGAGCCCGCTCTTCAACAACGTCGAAGCTACAGCGAAAGCTTCTCTGGAGTCGATCCAGA 347
Db 195 AGAGCCCGCTCTTCAACAACGTCGAAGCTACAGCGAAA-STTCTGTGGAGTCGATCCAGA 253

Qy 348 TTGTTATTAGAGAACTGAGGAAGAAAGGAACTCGAGTGGTGGATAGAGCAAGTCCA 407
Db 254 TTGTTATTAGAGAACTGAGGAAGAAAGGAACTCGAGTGGTGGATAGAGCAAGTCCA 313

Qy 408 GCTTCTCATCATGTGGCGGAGCCAGAGAATGGGGAACCTCATCTATCATGTGGTTT 467
Db 314 GCTTCTCATCATGTGGCGGAGG-CAGAGAATGGGGAACCTCATCTATCATGTGGTTT 372

Qy 468 CCAGGAGTGGC 478
Db 373 CCAGGAGTGGC 383

RESULT 11
US-09-621-976-15509
; Sequence 15509, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
```

```
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15509
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15509

Query Match 11.9%; Score 309; DB 4; Length 360;
Best Local Similarity 96.4%; Pred. No. 5.2e-64;
Matches 325; Conservative 10; Mismatches 0; Indels 2; Gaps 2;

Qy 109 CTGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCCAACCCCTT 168
Db 26 CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCCAACCCCTT 85

Qy 169 CTTTACGTTTACAACCGAATGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGCT 228
Db 86 CTTTACGTTTACAACCGAATGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGCT 145

Qy 229 GGTCTGTCTTCTTCGCGCCCTGCACAAACAGTCCAGCATGACGGTGTGGAAGCTCAGGA 288
Db 146 GGTCTGTCTTCTTCGCGCCCTGCACAAACAGTCCAGCATGACGGTGTGGAAGCTCAGGA 205

Qy 289 GAGCCCGCTCTTCAACAACGTCAGCTACAGCGAAAGCTTCTCTGGAGTCGATCCAGAT 348
Db 206 GAGCCCGCTCTTCAACAACGTCAGCTACAGCGAAA-STTCTGTGGAGTCGATCCAGAT 264

Qy 349 TGTATTAGAGAACTGAGGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAG 408
Db 265 TGTATTAGAGAACTGAGGAAGAAAGGAACTCGARTGGTGGATAAGAGCAAGTCCAG 324

Qy 409 CTTCTCTGATCATGTGGCGGAGCCAGAGAAATGGGGG 445
Db 325 CTTCTCTGATCATGTGGCGGARG-CAGAGAATGGGGG 360

RESULT 12
US-09-702-705-1380/c
; Sequence 1380, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 1380
; LENGTH: 269
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-1380

Query Match 10.3%; Score 269; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.6e-54;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 209 CTGGCCGCTGTGCTCGCTCGCTGTGCTTCTTGGCCGCTGCAACAACGATCCAGCATG 268
```



```
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-1380

Query Match      10.3%; Score 269; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.6e-54;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 CTGGCCGCTGGTCTCGTCTGCTGCTCTTCTGCGCGCTGCACAAACAGTCCAGCATG 268
Db 269 CTGGCCGCTGGTCTCGTCTGCTGCTCTTCTGCGCGCTGCACAAACAGTCCAGCATG 210

QY 269 ACGGTGATGGAGCTCAGGAGAGCCGCTCTTCAACACGCTACAGCGAAAGCTT 328
Db 209 ACGGTGATGGAGCTCAGGAGAGCCGCTCTTCAACACGCTACAGCGAAAGCTT 150

QY 329 CCTGTGGAGTCGATCCAGATTGTTAGAGGAACTGAGGAGAAAGGAAACCTCGAGTGG 388
Db 149 CCTGTGGAGTCGATCCAGATTGTTAGAGGAACTGAGGAGAAAGGAAACCTCGAGTGG 90

QY 389 TTGGATAGAGCAAGTCCAGCTTCTGATCATGTGGCGGAGCCAGAGAGATGGGGGAAA 448
Db 89 TTGGATAGAGCAAGTCCAGCTTCTGATCATGTGGCGGAGCCAGAGAGATGGGGGAAA 30

QY 449 CTCATCTATCAGTGGGTTTCCAGGAGTGG 477
Db 29 CTCATCTATCAGTGGGTTTCCAGGAGTGG 1
```

Search completed: June 29, 2005, 05:27:54
Job time : 299.707 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2005, 01:06:08 ; Search time 1019 Seconds
(without alignments)
15978.923 Million cell updates/sec

Title: US-08-731-499-4
Perfect score: 2605
Sequence: 1 CRAAGTCGAATTAACCTC.....CTGGCGTTACCACTTAAT 2605

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2605	100.0	2605	8	US-08-731-499-4
2	1076.2	41.3	1243	15	Sequence 4, Appl
3	1076.2	41.3	1243	17	Sequence 930, App
4	1073	41.2	1087	10	Sequence 930, App
5	603.6	23.2	634	11	Sequence 25, Appl
6	515.8	19.8	559	11	Sequence 727, App
7	464.4	17.8	507	11	Sequence 636, App
					Sequence 533, App

8	395	15.2	609	11	US-09-969-034-610
9	353.6	13.6	545	17	Sequence 610, App
10	317.4	12.2	349	11	Sequence 13, Appl
11	269	10.3	269	9	Sequence 665, App
12	269	10.3	269	9	Sequence 2264, App
13	269	10.3	269	9	Sequence 2400, App
14	269	10.3	269	9	Sequence 1380, App
15	269	10.3	269	9	Sequence 1380, App
16	269	10.3	269	9	Sequence 1380, App
17	269	10.3	269	9	Sequence 6790, App
18	269	10.3	269	14	Sequence 1380, App
19	269	10.3	269	16	Sequence 6790, App
20	269	10.3	269	16	Sequence 1380, App
21	269	10.3	269	17	Sequence 1380, App
22	269	10.3	269	17	Sequence 6790, App
23	269	10.3	269	19	Sequence 6790, App
24	267.4	10.3	269	9	Sequence 6790, App
25	267.4	10.3	269	14	Sequence 6647, App
26	267.4	10.3	269	17	Sequence 6647, App
27	267.4	10.3	269	17	Sequence 8647, App
28	267.4	10.3	269	19	Sequence 8647, App
29	260	10.0	260	9	Sequence 218, App
30	260	10.0	260	9	Sequence 3665, App
31	260	10.0	260	14	Sequence 3665, App
32	260	10.0	260	17	Sequence 3665, App
33	260	10.0	260	17	Sequence 3665, App
34	260	10.0	260	19	Sequence 3665, App
35	257	9.9	260	13	Sequence 1797, App
36	257	9.9	260	16	Sequence 1797, App
37	227	8.7	227	9	Sequence 7056, App
38	227	8.7	227	14	Sequence 7056, App
39	227	8.7	227	17	Sequence 7056, App
40	227	8.7	227	17	Sequence 7056, App
41	227	8.7	227	19	Sequence 7056, App
42	227	8.7	241	21	Sequence 611, App
43	219.2	8.4	469	20	Sequence 37999, A
44	218.2	8.4	335	20	Sequence 25157, A
45	169.4	6.5	217	10	Sequence 19060, A

ALIGNMENTS

RESULT 1
US-08-731-499-4
; Sequence 4, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOMBEL, David
; APPLICANT: ROMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435

1741	Db	CTATTGGAGACACATATGCTGGACTAAAGCAATAATATTTATTCTCAATGTCGTGCG	1801
1801	Qy	TAAACCTCAATGACTTAGAATGCTTTGCTATATATTTGCTCTCATGSCCTCAACCACTGGC	1860
1801	Db	TAAACCTCAATGACTTAGAATGCTTTGCTATATATTTGCTCTCATGSCCTCAACCACTGGC	1860
1861	Qy	TTTCTTTTJAGCTCTTGAACCAAGCCAACTGCTTCCGCTCAGGACAGAGATATTTTGGGA	1920
1861	Db	TTTCTTTTJAGCTCTTGAACCAAGCCAACTGCTTCCGCTCAGGACAGAGATATTTTGGGA	1920
1921	Qy	CTTCTCTTAAGAATTTCTATTTTCCTTAAATCTTTATCTGGGTAACTTAGTTTATTCCAACA	1980
1921	Db	CTTCTCTTAAGAATTTCTATTTTCCTTAAATCTTTATCTGGGTAACTTAGTTTATTCCAACA	1980
1981	Qy	CTTCAGATCCCTGGCTAAAAAATCTTCTTATAGAGCCTGTCTATGACACATGTCTCTCTTC	2040
1981	Db	CTTCAGATCCCTGGCTAAAAAATCTTCTTATAGAGCCTGTCTATGACACATGTCTCTCTTC	2040
2041	Qy	TCCAAACATCTCACCAGCACACATGTAGACTTAGAATTAGAACCTCTCTTTTCTTTTCA	2100
2041	Db	TCCAAACATCTCACCAGCACACATGTAGACTTAGAATTAGAACCTCTCTTTTCTTTTCA	2100
2101	Qy	TACTTTTCTCTATCATGCTTCCCTCCATATAATATTTTATATGTGTGAATGTCGTG	2160
2101	Db	TACTTTTCTCTATCATGCTTCCCTCCATATAATATTTTATATGTGTGAATGTCGTG	2160
2161	Qy	CCCCAAGTCAGTTTCCCTCACACTAACTATAAATCCGTAAAGCTGGGATCCTTCCAAATTTT	2220
2161	Db	CCCCAAGTCAGTTTCCCTCACACTAAACTATAAATCCGTAAAGCTGGGATCCTTCCAAATTTT	2220
2221	Qy	GATCACCACCTTAGTACAGTAGTAGAACACAGATAAAGATTTCAATTTGATTTTGTGGAATGAAT	2280
2221	Db	GATCACCACCTTAGTACAGTAGTAGAACACAGATAAAGATTTCAATTTGATTTTGTGGAATGAAT	2280
2281	Qy	GAATGAATTTGTTTTGTAGTAAAGTCTGGGGGAAACCCAGGTGAAGAGCCTTAGAAGCA	2340
2281	Db	GAATGAATTTGTTTTGTAGTAAAGTCTGGGGGAAACCCAGGTGAAGAGCCTTAGAAGCA	2340
2341	Qy	GGTTCGAATCCAAGCTAGATAGACTTAGTGTTACTCAAGAAAGGGTAGCCTGAAAAATAA	2400
2341	Db	GGTTCGAATCCAAGCTAGATAGACTTAGTGTTACTCAAGAAAGGGTAGCCTGAAAAATAA	2400
2401	Qy	GGTTCAAATTTATAGTCAAGAATAGTCAAGACATGGGCAAGAGAGTGTGCTCGTGCC	2460
2401	Db	GGTTCAAATTTATAGTCAAGAATAGTCAAGACATGGGCAAGAGAGTGTGCTCGTGCC	2460
2461	Qy	GAATTCGATATCAAGCTTATCGATACCGTTCGACCTCGAGGGGGGGCCCGGTACCCAAATTC	2520
2461	Db	GAATTCGATATCAAGCTTATCGATACCGTTCGACCTCGAGGGGGGGCCCGGTACCCAAATTC	2520
2521	Qy	GCCTATAGTGAGTCTGPTATTAACAATTCACCTGGCCGTGTTTTTACAACGTCGTGACTGGGA	2580
2521	Db	GCCTATAGTGAGTCTGPTATTAACAATTCACCTGGCCGTGTTTTTACAACGTCGTGACTGGGA	2580
2581	Qy	AAACCTCGCGGTACCCAACTTAAT	2605
2581	Db	AAACCTCGCGGTACCCAACTTAAT	2605

RESULT 2

```

US-10-037-270-930
; Sequence 930, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.

```


[illegible]


```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 577_
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-727

Query Match
Best Local Similarity 23.2%; Score 603.6; DB 11; Length 634;
Matches 629; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 321 GAAAGCTTCTGTGAGTCGATCCAGATTGTATTAGAGAACTGAGGAAGAAAGGAAC 380
Db 634 GAAAGCTTCTGTGAGTCGATCCAGATTGTATTAGAGAA--TTAAGAAGAAGGAACN 577

QY 381 TCGAGTGGTTGGATAAGAGCAAGTCAGCTTCTCTGATCATGTGGGGAGGCCAGAGAAT 440
Db 576 TCGAGTGGTTGGATAAGAGCAAGTCAGCTTCTCTGATCATGTGGGGAGGCCAGAGAAT 517

QY 441 GGGGAAACTCATCTATCAGTGGGTTTCCAGAGTGGCCAGAACTCCGTTCTTTACCC 500
Db 516 GGGGAAACTCATCTATCAGTGGGTTTCCAGAGTGG--CAGAACTCCGTTCTTTACCC 458

QY 501 TGTATGAATGACTAATGGGGAAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAG 560
Db 457 TGTATGAATGACTAATGGGGAAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAG 398

QY 561 CCACCTCTACTGGGGCTCTGAGGCCCTACAGCAGGAGCAGAGGCCAGATCATCACTG 620
Db 397 CCACCTCTACTGGGGCTCTGAGGCCCTACAGCAGGAGCAGAGGCCAGATCATCACTG 338

QY 621 TCAGCGATGGCGAGGCGTCAAGTTCTTCTAGCAGGACCTGTCTCTCTTTACTTCTTAC 680
Db 337 TCAGCGATGGCGAGGCGTCAAGTTCTTCTAGCAGGACCTGTCTCTCTTTACTTCTTAC 278

QY 681 CTCGCCACCTTTCCAGGGCTTTCAAAGGAGACAGACCAGTGTTCCCCCAAAGACTGGATC 740
Db 277 CTCGCCACCTTTCCAGGGCTTTCAAAGGAGACAGACCAGTGTTCCCCCAAAGACTGGATC 218

QY 741 TGTGACTCCAGACACTCAAAGGACTCCAGTCTCGAAGGCTGGAGCTGGGGATGGGTT 800
Db 217 TGTGACTCCAGACACTCAAAGGACTCCAGTCTCGAAGGCTGGAGCTGGGGATGGGTT 158

QY 801 TCTCACACCCATATGTCTGCTCCCTTGGATAGGTTGAGGCTGAAGCACAGGAGAAAT 860
Db 157 TCTCACACCCATATGTCTGCTCCCTTGGATAGGTTGAGGCTGAAGCACAGGAGAAAT 98

QY 861 ATGTGCTTTCTTCTCGCCCTTACTCTCTTTCCCATCTAGACTGTCTTGGCCAGGGTCTG 920
Db 97 ATGTGCTTTCTTCTCGCCCTTACTCTCTTTCCCATCTAGACTGTCTTGGCCAGGGTCTG 38

QY 921 TAAACCTGACACTTTATATGTGTTCCACATGTAAGT 957
Db 37 TAAACCTGGCACTTTATATGTGTTCCACATGTAAGT 1

RESULT 6
US-09-969-034-636
; Sequence 636, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
```

```
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 636
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-034-636

Query Match
Best Local Similarity 19.8%; Score 515.8; DB 11; Length 559;
Matches 517; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 109 CTGGGCTACTAGTAGGGGATGAGTTTCGAGTGGCCGTCGAGTATCCGTTCCCAACCTT 168
Db 16 CTGGGCTACTAGTAGGGGATGAGTTTCGAGTGGCCGTCGAGTATCCGTTCCCAACCTT 75

QY 169 CTTTACGTTTACAACCGAATGTGGACACTCGGCAGAGCAGCTGGCCGCTGGTGTCTCGCT 228
Db 76 CTTTACGTTTACAACCGAATGTGGACACTCGGCAGAGCAGCTGGCCGCTGGTGTCTCGCT 135

QY 229 GGTCTCTGTCTTCTGCGCCCTGCAAAACAGTCCAGCATGACGGTGTATGGAAGCTCAGGA 288
Db 136 GGTCTCTGTCTTCTGCGCCCTGCAAAACAGTCCAGCATGACGGTGTATGGAAGCTCAGGA 195

QY 289 GAGCCGCTCTTCAACACGTCACAGTACAGGAAAGCTTCTGTGGAGTCGATCCAGAT 348
Db 196 GAGCCGCTCTTCAACACGTCACAGTACAGGAAAGCTTCTGTGGAGTCGATCCAGAT 255

QY 349 TGTATTACAGGAACTGAGGAAGAAAGGAACTCCAGTGGTTGGATAAGAGCAAGTCCAG 408
Db 256 TGTATTACAGGAACTGAGGAAGAAAGGAACTCCAGTGGTTGGATAAGAGCAAGTCCAG 315

QY 409 CTTCTCTGATCATGTGGCGGAGCCAGAAAGAAATGCGGGAACCTCATCTATCATGTGGTTTC 468
Db 316 CTTCTCTGATCATGTGGCGGAGCCAGAAAGAAATGCGGGAACCTCATCTATCATGTGGTTTC 375

QY 469 CAGGAGTGGCCAGAACAACTCGTCTTTTACCTGTATGACTGACTAATGGGGAAGACAC 528
Db 376 CAGGAGTGGCCAGAACAACTCGTCTTTTACCTGTATGACTGACTAATGGGGAAGACAC 435

QY 529 AGAGGATGAGGAGTTCCACGGCTGGATGAAGCCACTTACTTGGGGGCTCTGACAGGCCCT 588
Db 436 AGAGGATGAGGAGTTCCACGGCTGGATGAAGCCACTTACTTGGGGGCTCTGACAGGCCCT 495

QY 589 ACAGCAGAGCACAAAGCCGAGATCATCTGTGACCGA 627
Db 496 ACAGCAGAGCACAAAGCCGAGATCATCTGTGACCGA 534

RESULT 7
US-09-969-034-533/c
; Sequence 533, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 533
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 327, 340, 357, 368, 469, 498
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-533

Query Match 17.8%; Score 464.4; DB 11; Length 507;
Best Local Similarity 97.0%; Pred. No. 1.9e-96;
Matches 491; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

QY 455 TATCAGTGGGTTTCCAGGA--GTGGCCAGAACCACTCCGTCTTTACCCCTGTATGAAC-TG 511
DB 506 TATCAGTGGTGTTCAGGAAGTGGCCAGAACCAACTCNGTCTTTTACCCCTGTATGAACCTG 447

QY 512 ACTAATGGGGAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAGCACTCTACTG 571
DB 446 ACTAATGGGGAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAGCACTCTTACTG 387

QY 572 CGGGCTCTGCAGGCCCTACAGAGGAGCAAGGCCGAGATCATCACTGTACGCGATGCG 631
DB 386 CGGGCTCTGCAGGCCCTTANAGCAGGAGCANAGGCCGAGATCATCACTGTACGCGATGCG 327

QY 632 CGAGGCGTCAAGTCTTCTAGCAGGACCTGTCTCCCTTTTACTTCTTACCTTCCACCTTT 691
DB 326 CGAGGCGTCAAGTCTTCTAGCAGGACCTGTCTCCCTTTTACTTCTTACCTTCCACCTTT 267

QY 692 CGAGGCGTTCAGAGGAGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTCCAC 751
DB 266 CGAGGCGTTCAGAGGAGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTCCAC 207

QY 752 CAGACTCAAAAGACTCCAGTCTGAAGGCTGGGACTGGGGATGGTTTCTCACACCCC 811
DB 206 CAGACTCAAAAGACTCCAGTCTGAAGGCTGGGACTGGGGATGGTTTCTCACACCCC 147

QY 812 ATATGTCTCTCCCTGGATAGGCTGAGGCTGAAGCACACGAGGAGAAAATATGTCTTCTT 871
DB 146 ATATGTCTCTCCCTGGATAGGCTGAGGCTGAAGCACACGAGGAGAAAATATGTCTTCTT 87

QY 872 CTCGCCCTACCTCTTTCCTTCCATCTTATAGCTGTCTTTCAGCCAGGGTCTGTAACTGACA 931
DB 86 CTCGCCCTACCTCTTTCCTTCCATCTTATAGCTGTCTTTCAGCCAGGGTCTGTAACTGACA 27

QY 932 CTTTATATGTCTTACACATGTAAGT 957
DB 26 CTTTATATGTCTTACACATGTAAGT 1

RESULT 8
US-09-969-034-610
; Sequence 610, Application US/09969034
; Publication No. US2004011068A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 610
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 468, 478, 491, 501, 505, 574, 576, 577, 604
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-610

Query Match 15.2%; Score 395; DB 11; Length 609;
Best Local Similarity 96.5%; Pred. No. 2e-80;
Matches 447; Conservative 0; Mismatches 10; Indels 6; Gaps 4;

QY 112 GGCTACTACGATGGCGGATGAGTTTCAGTGGCGGTGGCAGTATCGCTTCCACCCCTTCTT 171
DB 5 GGGGACTACGATGGTGTGATGAGTTTCAGTGGCGGTGGCAGTATCGCTTCCACCCCTTCTT 64

QY 172 TAGCTTACCAACCGAATGTGGACACTCGGCAGAACGAGCTGGCGCCTGTGTCTCGTGTGT 231
DB 65 TAGCTTACCAACCGAATGTGGACACTCGGCAGAACGAGCTGGCGCCTGTGTCTCGTGTGT 124

QY 232 CTGTCTCTTCTCCGCTCGCACAAACAGTCCAGCATGACGGTATGGAAGCTTCCAGGAG 291
DB 125 CTGTCTCTTCTCCGCTCGCACAAACAGTCCAGCATGACGGTATGGAAGCTTCCAGGAG 184

QY 292 CCGCTCTTTCACCAACGCTCAAGCTACAGCGAAAGCTTCTGTGGAGTCCGATCCAGATTGT 351
DB 185 CCGCTCTTTCACCAACGCTCAAGCTACAGCGAAAGCTTCTGTGGAGTCCGATCCAGATTGT 244

QY 352 ATTAGAGGAATCGAGGAAGAAAGGAACTCTCGAGTGGTGGATTAAGACAAAGTCCAGCTT 411
DB 245 ATTAGAGGAATCGAGGAAGAAAGGAACTCTCGAGTGGTGGATTAAGACAAAGTCCAGCTT 304

QY 412 CCGTATCATGTGGCGGAGCCAGAGAAATGGGGGAACTCATCTATCAGTGGTGTCCAG 471
DB 305 CCGTATCATGTGGCGGAGCCAGAGAAATGGGGGAACTCATCTATCAGTGGTGTCCAG 361

QY 472 GAGTGGCCAGAACTCACTCGTCTTTACCTGTATGAACCTGACTTAATGGGGAAGACACAGA 531
DB 362 GAGTGGC--AGAACACTCGTCTTTACCTGTATGAACCTGACTTAATGGGGAAGACACAGA 419

QY 532 GGATGAGGAGTTCCACGGGCTGGATGAAGCCACTCTACTGCGG 574
DB 420 GGATGAGGAGTTCCACGGGCTGGATGAAGCCACTCTACTGCGG 461

RESULT 9
US-10-115-482-13
; Sequence 13, Application US/10115482
; Publication No. US20030212257A1
; GENERAL INFORMATION:
; APPLICANT: Spytex, et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS
; TITLE OF INVENTION: OF USING THE SAME
; FILE REFERENCE: 21404-322D
; CURRENT APPLICATION NUMBER: US/10/115,482
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/285,325

;; PRIOR FILING DATE: 2001-04-19
;; PRIOR APPLICATION NUMBER: 60/285,890
;; PRIOR FILING DATE: 2001-04-23
;; PRIOR APPLICATION NUMBER: 60/286,068
;; PRIOR FILING DATE: 2001-04-24
;; PRIOR APPLICATION NUMBER: 60/286,292
;; PRIOR FILING DATE: 2001-04-25
;; PRIOR APPLICATION NUMBER: 60/287,213
;; PRIOR FILING DATE: 2001-04-27
;; PRIOR APPLICATION NUMBER: 60/288,257
;; PRIOR FILING DATE: 2001-05-02
;; PRIOR APPLICATION NUMBER: 60/291,134
;; PRIOR FILING DATE: 2001-05-15
;; PRIOR APPLICATION NUMBER: 60/282,020
;; PRIOR FILING DATE: 2001-04-06
;; PRIOR APPLICATION NUMBER: 60/291,725
;; PRIOR FILING DATE: 2001-05-17
;; PRIOR APPLICATION NUMBER: 60/294,771
;; PRIOR FILING DATE: 2001-05-31
;; PRIOR APPLICATION NUMBER: 60/296,965
;; PRIOR FILING DATE: 2001-06-08
;; PRIOR APPLICATION NUMBER: 60/299,128
;; PRIOR FILING DATE: 2001-06-08
;; NUMBER OF SEQ ID NOS: 149
;; SEQ ID NO 13
;; LENGTH: 545
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (18)..(534)
US-10-115-482-13

Query Match 13.6%; Score 353.6; DB 17; Length 545;
Best Local Similarity 84.6%; Pred. No. 6.4e-71;
Matches 435; Conservative 0; Mismatches 69; Indels 10; Gaps 3;

QY 109 CTGGCTACTACGATGGCGATGAGTTTCAGTGGCGCGTGGCAGTATCGCTTCCACCCCTT 168
Db 8 CTGGCTATTATGATGGTGAGCAATTCAGTGGTGTGGCAGTATCGCTTCCGCC--- 64
QY 169 CTTTACGTTTACAAACCGAATGTGACACTCGGCAGAGCAGCTGGCGGCTGTGTCTCGCT 228
Db 65 CTTTACATTACAGCTGAAGTGGCCACTTGGCAGAGCAGCTGGCCACCTGGTGTGT 124
QY 229 GGTCTGTCTTCTCCGCGCTGCACAAACAGTCCAGCATGACGGTGATGGAAGCTCAGGA 288
Db 125 GGTCTGTCTCATCTGCTGCTGCACAGACAGTCAAGCATGATGGTTATGGATGCTCAGGA 184
QY 289 GAGCCCGCTCTTCAACAACGCTCAAGCTACAGGAAAGCTTCTGTGGAGTGCATCCAGAT 348
Db 185 GATCTGTCTTTCAGCAACATCAAGCT---GTGGAGCTTCTGTGGGATCATCCAGGT 241
QY 349 TGTATTAGAGGAACGTGAGGAAGAAAGGAAACCTCGAGTGGTGGTGGATAAGAGCAAGTCCAG 408
Db 242 TGTATTAGAGGAACGTGAGGAAGAAAGGAACTTACAGTGGCTGGGATAAGAGCAAGTCTAG 301
QY 409 CTTCTGTATCATGTGGCGGAGCCAGAGAATGGGGGAAACCTCATCTATCATGTGGGTTTC 468
Db 302 TTTCTTAATCATGTGGCGGAGCCAGAGAATGGGGGAAACCTCATCTATCATGTGGGTTTC 361
QY 469 CAGGAGTGGCCAGAACAACTCCGTCTTTACCTGTATGAACTGACTTAATGGGGAAGACAC 528
Db 362 CAGGAGTGGCCAGAACAACTCCGTACTTACCTGTATGAGCTGACCAATGGGGAAGACAT 421
QY 529 AGAGGATGAGGAGTTCCACGGGCTGGATGAAGCCACTCTACTCGGGCTCTGACGGCCCT 588
Db 422 AGAGAATGAGGTGTTCCACGGACTAAAGGAGGC---CTTCTGTGGGCTCTGACGGCCCT 477
QY 589 ACAGCAGAGCAACAGCGCGAGATCATCTGTC 622
Db 478 TCAGTAGGAACATAGGCTGAGATCATCACCATC 511

RESULT 10
US-09-969-034-665
;; Sequence 665, Application US/09969034
;; Publication No. US20040110668A1
;; GENERAL INFORMATION:
;; APPLICANT: Burgess, Christopher C.
;; APPLICANT: Astle, Jon H.
;; APPLICANT: Carroll, Eddie III
;; APPLICANT: Catino, Theodore J.
;; APPLICANT: Dwivedi, Poorima
;; APPLICANT: Molino, Gary A.
;; APPLICANT: Thiagalingam, Arunthathi
;; APPLICANT: Lewis, Marcia E.
;; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
;; TITLE OF INVENTION: Expressed in Cancer Tissue
;; FILE REFERENCE: 1657/1032
;; CURRENT APPLICATION NUMBER: US/09/969,034
;; CURRENT FILING DATE: 2001-10-02
;; PRIOR APPLICATION NUMBER: 60/237,271
;; PRIOR FILING DATE: 2000-02-10
;; NUMBER OF SEQ ID NOS: 4494
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 665
;; LENGTH: 349
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 233, 274, 278, 321
;; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-665

Query Match 12.2%; Score 317.4; DB 11; Length 349;
Best Local Similarity 98.5%; Pred. No. 1.1e-62;
Matches 318; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 637 CGTCAAGTTCTTTAGCAGGAGCCTGCTCCCTTTTACTTCTTACCTCCACCTTTCCAGG 696
Db 1 CGTCAAGTTCTTTAGCAGGAGCCTGCTCCCTTTTACTTCTTACCTCCACCTTTCCAGG 60
QY 697 GCTTTCAAAGGAGACAGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACAGAC 756
Db 61 GCTTTCAAAGGAGACAGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACAGAC 120
QY 757 TCAAAAGACTCCAGTCTCTGAAGGCTGGGACTGGGGATGGTTTCTCACACCCCATATG 816
Db 121 TCAAAAGACTCCAGTCTCTGAAGGCTGGGACTGGGGATGGTTTCTCACACCCCATATG 180
QY 817 TCTGTCCCTTCGATAGGTGAGGCTGAAGCACCAGGGAGAAAATATGTGCTTCTTCTCGC 876
Db 181 TCTGTCCCTTCGATAGGTGAGGCTGAAGCACCAGGGAGAAAATATGTGCTTCTTCTCGC 240
QY 877 CTTACTCTCTTTCCCATCTTAGACTGTCTTTGAGCCAGGGTCTGTAAACCTGACACTTTA 936
Db 241 CTTACTCTCTTTCCCATCTTAGACTGTCTTTGANCANGGTCTGTAAACCTGACACTTTA 300
QY 937 TATGTGTTACACATGTAAGTAC 959
Db 301 TATGTGTTACACATGTAAGNCC 323

RESULT 11
US-09-998-598-2264/c
;; Sequence 2264, Application US/09998598
;; Patent No. US20020150922A1
;; GENERAL INFORMATION:
;; APPLICANT: Stolk, John A.
;; APPLICANT: Xu, Jiangchun
;; APPLICANT: Chenault, Ruth A.
;; APPLICANT: Mesgher, Madelein Joy
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 15:06:07 ; Search time 5606.03 Seconds
(without alignments)
17687.639 Million cell updates/sec

Title: US-08-731-499-4
Perfect score: 2605
Sequence: 1 CAAGCTCGAATTAACCTC.....CTGGCGTTACCACTTAAT 2605

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gssi:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1041	40.0	1041	3	CR620674	full-leng
2	1040	39.9	1060	3	CR611687	full-leng
3	1016	39.0	1016	3	CR615047	full-leng
4	965.4	37.1	1021	1	AL561088	AL561088
5	947.2	36.4	1079	5	BX357325	BX357325
6	923	35.4	1039	5	BM927839	AGENCOURT
7	914.8	35.1	971	1	AL582492	AL582492
8	903	34.7	1005	4	BM554015	AGENCOURT
9	889	34.1	982	6	CA487570	AGENCOURT
10	879.2	33.8	972	5	BX383031	BX383031
11	868	33.3	918	5	BQ672154	ILLUMIGEN
12	859	33.0	936	7	COS81180	ILLUMIGEN
13	854.6	32.8	1049	5	BM919274	AGENCOURT
14	844.8	32.4	925	4	BG675348	602621657
15	843.8	32.4	990	4	BM555365	AGENCOURT
16	838	32.2	959	5	BX357450	BX357450
17	831.6	31.9	1046	4	BM810698	AGENCOURT
18	830.4	31.9	901	4	BM018273	603645927
19	826.4	31.7	961	5	BQ882132	AGENCOURT
20	822.8	31.6	897	6	CA455057	AGENCOURT
21	814.6	31.3	859	4	BG674662	602620732
22	813.6	31.2	919	4	BG742758	602622982
23	811.4	31.1	957	7	COT26143	ILLUMIGEN
24	811.2	31.1	943	5	BQ673129	AGENCOURT

25	811	31.1	909	7	COS82238	ILLUMIGEN
26	807	31.0	908	5	BU540421	AGENCOURT
27	804.8	30.9	877	5	BU168204	AGENCOURT
28	799.4	30.7	898	6	CA489194	AGENCOURT
29	798.8	30.7	981	4	BM474026	AGENCOURT
30	796.6	30.6	877	4	BI550414	603192475
31	790.8	30.4	842	1	AL528911	AL528911
32	790.6	30.3	928	6	CA489833	AGENCOURT
33	778.6	29.9	923	5	BQ646055	AGENCOURT
34	776.8	29.8	779	1	AL528912	AL528912
35	775.2	29.8	863	5	BU183716	AGENCOURT
36	775.2	29.8	877	5	BU153711	AGENCOURT
37	772.4	29.7	911	6	CA487561	AGENCOURT
38	771.2	29.6	1027	5	BQ053239	AGENCOURT
39	770.4	29.6	1074	3	AK003715	Mus muscu
40	769.2	29.5	1057	4	BM476382	AGENCOURT
41	767.8	29.5	872	6	CD388777	AGENCOURT
42	766.8	29.4	1067	3	AK002500	Mus muscu
43	751	28.8	974	4	BM561840	AGENCOURT
44	750.8	28.8	829	4	BG281239	602402192
45	745.4	28.6	803	5	BQ773588	UI-H-PEO-

ALIGNMENTS

RESULT 1
CR620674
LOCUS CR620674 1041 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DD004YG22 of Neuroblastoma Cot
50-normalized of Homo sapiens (human).
ACCESSION CR620674.1 GI:50501481
VERSION HTC; CNSLT cDNA.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1041)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1041)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by life technologies, a
division of invitrogen.
FEATURES
source
1..1041
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD004YG22"
/tissue_type="Neuroblastoma Cot 50-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 40.0%; Score 1041; DB 3; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.4e-217;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 129 TGAGTTTCGAGTGGCGGTGAGTATCGTTCACCCCTTCTTACGTTACACCGAATG 188
DB 1 TGAGTTTCGAGTGGCGGTGAGTATCGTTCACCCCTTCTTACGTTACACCGAATG 60


```
QY 589 ACAGCAGGAGCACAAGGCCGAGATCATCTCTCAGCGATGCGCGAGGCGTCAAGTTCTT 648
Db 501 ACAGCAGGAGCACAAGGCCGAGATCATCTCTCAGCGATGCGCGAGGCGTCAAGTTCTT 560
QY 649 CTAGCAGGAGCTGTCTCCCTTTACTTCTTACTCTCCACCTTTCCAGGCGCTTTCAAAGG 708
Db 561 CTAGCAGGAGCTGTCTCCCTTTACTTCTTACTCTCCACCTTTCCAGGCGCTTTCAAAGG 620
QY 709 AGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTTCCACAGACTCAAAAGGACTC 768
Db 621 AGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTTCCACAGACTCAAAAGGACTC 680
QY 769 CAGTCTGAAGGCTGGGACCTGGGATGGGTTTCTCTCACCCCATATGTCTGCTCCCTGG 828
Db 681 CAGTCTGAAGGCTGGGACCTGGGATGGGTTTCTCTCACCCCATATGTCTGCTCCCTGG 740
QY 829 ATAGGCTGAGGCTGAAGCACAAGGAGGAAATATGTCTTCTTCGCGCTACCTCCCTT 888
Db 741 ATAGGCTGAGGCTGAAGCACAAGGAGGAAATATGTCTTCTTCTCGCGCTACCTCCCTT 800
QY 889 CCCATCTAGACTGTCTTGAGCCAGGCTGTGAACCTGACACTTTATATGTGTTCACA 948
Db 801 CCCATCTAGACTGTCTTGAGCCAGGCTGTGAACCTGACACTTTATATGTGTTCACA 860
QY 949 CATGTAAGTACATACACATGCGGCTGAGACACATGTTCTGTCTCTCTCTCCAC 1008
Db 861 CATGTAAGTACATACACATGCGGCTGAGACACATGTTCTGTCTCTCTCTCCAC 920
QY 1009 CCCTTTAGCTGTGTGCTCCCTTCTCAGGCTGTGTGATCTTCTTCTAGGGGATGG 1068
Db 921 CCCTTTAGCTGTGTGCTCCCTTCTCAGGCTGTGTGATCTTCTTCTAGGGGATGG 980
QY 1069 GGAAGCCCTGGCTGAGGAGGCTTCCAGGCAATATGAAGATAGGAGGCCACGCGGCTG 1128
Db 981 GGAAGCCCTGGCTGAGGAGGCTTCCAGGCAATATGAAGATAGGAGGCCACGCGGCTG 1040
QY 1129 GCAGTGAGAGGTGTGGCCCC 1148
Db 1041 GCAGTGAGAGGTGTGGCCCC 1060

RESULT 3
LOCUS CR615047
DEFINITION full-length cDNA clone CS0DC007YC08 of Neuroblastoma Cot
25-normalized of Homo sapiens (human).
ACCESSION CR615047
VERSION CR615047.1 GI:50495854
KEYWORDS HTC; cDNA
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1016)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1016)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. .1016

FEATURES
source
```

ORIGIN

```
Query Match 39.0%; Score 1016; DB 3; Length 1016;
Best Local Similarity 100.0%; Pred. No. 4.2e-212;
Matches 1016; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 TGAGTTTCAGTGGCGGTGGCAGTATCGCTTCCACCCCTTCTTTACGTAGCAACGAATG 188
Db 1 TGAGTTTCAGTGGCGGTGGCAGTATCGCTTCCACCCCTTCTTTACGTAGCAACGAATG 60
QY 189 TGGACACTCGCAGCAAGCAGCTGGCGCTGTGTCTCGTGTCTCTGTCTCTTCTCCCGC 248
Db 61 TGGACACTCGCAGCAAGCAGCTGGCGCTGTGTCTCGTGTCTCTGTCTCTTCTCTCCCGC 120
QY 249 TGCACAAACAGTCCACGATGACGGTGTGAAGCTTCAGGAGAGCCCGCTCTTTCAACAACG 308
Db 121 TGCACAAACAGTCCACGATGACGGTGTGAAGCTTCAGGAGAGCCCGCTCTTTCAACAACG 180
QY 309 TCAAGCTACAGGAAAGCTTCTCTGTGGAGTGCATCCAGATTGTATTAGAGGAATGAGGA 368
Db 181 TCAAGCTACAGGAAAGCTTCTCTGTGGAGTGCATCCAGATTGTATTAGAGGAATGAGGA 240
QY 369 AGAAGGGAACCTCAGTGGTTGGATAGAGCAAGTCCAGCTTCTGTATCATGTGCGGA 428
Db 241 AGAAGGGAACCTCAGTGGTTGGATAGAGCAAGTCCAGCTTCTGTATCATGTGCGGA 300
QY 429 GGCACAGAAAGATGGGGGAAACTCATCTATCATGTGGTTTCCAGGAGTGCGCCAGAACAACT 488
Db 301 GGCACAGAAAGATGGGGGAAACTCATCTATCATGTGGTTTCCAGGAGTGCGCCAGAACAACT 360
QY 489 CCGTCTTTACCTGTATGAACTGACTAATGGGGAAGACACAGAGATGAGGATGTCACG 548
Db 361 CCGTCTTTACCTGTATGAACTGACTAATGGGGAAGACACAGAGATGAGGATGTCACG 420
QY 549 GCCTGGATGAAGCCACTCTACTCGGGCTCTGCAGGCCCTACAGCAGGAGCAAGGCGG 608
Db 421 GCCTGGATGAAGCCACTCTACTCGGGCTCTGCAGGCCCTACAGCAGGAGCAAGGCGG 480
QY 609 AGATCATCACTGTTCAGCGATGCGCGAGCGTCAAGTTCTTCTAGCAGGACCTGTCTCC 668
Db 481 AGATCATCACTGTTCAGCGATGCGCGAGCGTCAAGTTCTTCTAGCAGGACCTGTCTCC 540
QY 669 TTTACTTCTTACCTCCACCTTTTCCAGGGCTTTTCAAAGGAGACAGACCGAGTGTCGCC 728
Db 541 TTTACTTCTTACCTCCACCTTTTCCAGGGCTTTTCAAAGGAGACAGACCGAGTGTCGCC 600
QY 729 AAAGACTGGATCTGTGACTTCCACAGACTCAAAAGGACTCCAGTCTCTGAAGGCTGGGACC 788
Db 601 AAAGACTGGATCTGTGACTTCCACAGACTCAAAAGGACTCCAGTCTCTGAAGGCTGGGACC 660
QY 789 TGGGATGGGTTTCTCACACCCCATATGTCTGTCTCTTGGATAGGCTGAGGCTGAAGCAC 848
Db 661 TGGGATGGGTTTCTCACACCCCATATGTCTGTCTCTTGGATAGGCTGAGGCTGAAGCAC 720
QY 849 CAGGAGAAATATGTGCTTCTTCTGCGCCCTACCTCTCTTCCCATCTCTAGACTGCTCTTG 908
Db 721 CAGGAGAAATATGTGCTTCTTCTGCGCCCTACCTCTCTTCCCATCTCTAGACTGCTCTTG 780
QY 909 AGCCAGGCTGTAAACCTGACACTTTATATGTGTTTACATGTAAGTACATACACACA 968
Db 781 AGCCAGGCTGTAAACCTGACACTTTATATGTGTTTACATGTAAGTACATACACACA 840
QY 969 TGGCGCTTCAGACATGTTCTGTCTCTCTCTCTCCACCCCTTTAGCTGTGTGCT 1028
Db 841 TGGCGCTTCAGACATGTTCTGTCTCTCTCTCTCTCCACCCCTTTAGCTGTGTGCT 900
QY 1029 CCCTTCTCAGGCTGTGTGATCTTCTCTAGGGGATGGGGAGAGCCCTGGCTGAGGCA 1088
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC007YC08"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"
```

|||||
901 CCTTCTCAGGCTGGTGGATCTTCTTAGGAGTGGGGAAGCCCTGGCTGCAGGCA 960
|||||
1089 GCCTTCAGGCAATATGAAGATAGAGGCCACCGGCTGGCAGTGAGAGGTGTGG 1144
|||||
961 GCCTTCAGGCAATATGAAGATAGAGGCCACCGGCTGGCAGTGAGAGGTGTGG 1016
|||||

RESULT 4
AL561088 1021 bp mRNA linear EST 02-APR-2004
LOCUS
DEFINITION
AL561088 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0D001Y06 5-PRIME, mRNA sequence.
ACCESSION
AL561088
VERSION
AL561088.3 GI:46186449
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1021)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31285217.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: secre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8673.f

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?g=CS0D001Y06&g=8673.f>.

FEATURES

source

1..1021
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D001Y06"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 37.1%; Score 965.4; DB 1; Length 1021;
Best Local Similarity 98.8%; Pred. No. 5.4e-201;
Matches 977; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
Qy 113 GCTACTAGATGGCGATGAGTTTCGAGTGGCGGAGTATCGCTCCACCCCTCTTT 172
Db 29 GCTACTAGATGGCGATGAGTTTCGAGTGGCGGAGTATCGCTCCACCCCTCTTT 88
Qy 173 ACGTTTACACCGAATGTGACACTCGGCAGAGCAGCTGGCGCTGGTGGCTGGCTGTC 232
Db 89 ACGTTTACACCGAATGTGACACTCGGCAGAGCAGCTGGCGCTGGTGGCTGGCTGTC 148
Qy 233 CTGCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTGAATGAAGCTCAGGAGAGC 292
Db 149 CTGCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTGAATGAAGCTCAGGAGAGC 208
Qy 293 CCGCTCTTCAACAGCTCAAGCTACAGGAAAGCTTCTGTGGAGTCGATCCAGATTGTA 352
Db 209 CCGCTCTTCAACAGCTCAAGCTACAGGAAAGCTTCTGTGGAGTCGATCCAGATTGTA 268

Qy 353 TTAGAGGAACCTGAGAGAGAGGAACTCGAGTGGTTGGATAGAGCAAGTCCAGCTTC 412
Db 269 TTAGAGGAACCTGAGAGAGAGGAACTCGAGTGGTTGGATAGAGCAAGTCCAGCTTC 328
Qy 413 CTGATCATGTGCGGAGGCGCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTCACAG 472
Db 329 CTGATCATGTGCGGAGGCGCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTCACAG 388
Qy 473 AGTGGCCAGAACTCGCTCTTACCTGTATGAATGACTTAATGGGAGAGACAGAG 532
Db 389 AGTGGCCAGAACTCGCTCTTACCTGTATGAATGACTTAATGGGAGAGACAGAG 448
Qy 533 GATGAGAGTTTCCACGGGCTGATGAAGCACTCTACTGCGGGCTCTCAGGCCCTACAG 592
Db 449 GATGAGAGTTTCCACGGGCTGATGAAGCACTCTACTGCGGGCTCTCAGGCCCTACAG 508
Qy 593 CAGGAGCAAGGGCCGAGATCATCTGTTCAGCGATGGCCGAGGCGTCAAGTTCCTTCAG 652
Db 509 CAGGAGCAAGGGCCGAGATCATCTGTTCAGCGATGGCCGAGGCGTCAAGTTCCTTCAG 568
Qy 653 CAGGAGCTGTCTCCCTTACTTCTTACCTGCCACCTTTCAGGGCTTTCAAAAGGAGAC 712
Db 569 CAGGAGCTGTCTCCCTTACTTCTTACCTGCCACCTTTCAGGGCTTTCAAAAGGAGAC 628
Qy 713 AGACCCAGTGTCCCCCAAGAGCTGGATCTGTGACTCCACACAGACTCAAAAGGAGCTCCAGT 772
Db 629 AGACCCAGTGTCCCCCAAGAGCTGGATCTGTGACTCCACACAGACTCAAAAGGAGCTCCAGT 688
Qy 773 CCTGAAGCTGGGACCTGGGATGGTTCACACCCCATATGTCTCTCTCCCTTCCCTCCCA 832
Db 689 CCTGAAGCTGGGACCTGGGATGGTTCACACCCCATATGTCTCTCTCCCTTCCCTCCCA 748
Qy 833 GGTGAGCTGAAGCACCAGGAGAAATATGTGTCTTCTCGCCCTACCTCTCTTCCCA 892
Db 749 GGTGAGCTGAAGCACCAGGAGAAATATGTGTCTTCTCTCGCCCTACCTCTCTTCCCA 808
Qy 893 TCCTAGACTGTCTTGAGCAGGCTGTAAACCTGACACTTTATATGTGTTCACACATG 952
Db 809 TCCTAGACTGTCTTGAGCAGGCTGTAAACCTGACACTTTATATGTGTTCACACATG 868
Qy 953 TAAGTACATACACATGCGCTCGAGCAGATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1012
Db 869 TAAGTACATACACATGCGCTCGAGCAGATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 928
Qy 1013 TTAGTCTGTGTCT 1072
Db 929 TTAGTCTGTGTCT 988
Qy 1073 GCCCTGGCTGAGGAGGAGCTTCCAGGCAA 1101
Db 989 G-CCTGGSTGCAGSAGCYTTCAGGGA 1016
RESULT 5
BX357325/c 1079 bp mRNA linear EST 23-APR-2004
LOCUS
DEFINITION
BX357325 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1024YM13 3-PRIME, mRNA sequence.
ACCESSION
BX357325
VERSION
BX357325.2 GI:46551167
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1079)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30380160.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE

FEATURES		Location/Qualifiers	
source		1..1005	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5742295"	
		/tissue_type="medulla"	
		/lab_host="DH10B"	
		/clone_lib="NIH_MGC_119"	
		/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."	
ORIGIN			
Query Match		34.7%; Score 903; DB 4; Length 1005;	
Best Local Similarity		97.1%; Pred. No. 2.7e-187;	
Matches	973;	Conservative 0; Mismatches 21; Indels 8; Gaps 5;	
QY	114	CTACTACGATGCGATGATGTTTCGAGTGGCCGCTGGCAGTATCGTCTCCACCCCTCTTTA	173
Db	1	CTACTACGATGCGATGATGTTTCGAGTGGCCGCTGGCAGTATCGTCTCCACCCCTCTTTA	60
QY	174	CGTTACACCGAATGTGACACTCGGCAGAGCAGCTGGCCGCTGGTCTCGCTGGTCC	233
Db	61	CGTTACACCGAATGTGACACTCGGCAGAGCAGCTGGCCGCTGGTCTCGCTGGTCC	120
QY	234	TGTCCTTCTGCGCCCTGCACAAACAGTCAGCATGACGGTGATGGAAGCTCAGGAGAGCC	293
Db	121	TGTCCTTCTGCGCCCTGCACAAACAGTCAGCATGACGGTGATGGAAGCTCAGGAGAGCC	180
QY	294	CGCTCTTCAACAACTCAAGCTACAGCAAAAGCTCTCTGTGGAGTCGATCCAGATTGTAT	353
Db	181	CGCTCTTCAACAACTCAAGCTACAGCAAAAGCTCTCTGTGGAGTCGATCCAGATTGTAT	240
QY	354	TAGAGGAACTGAGGAGAAAGGGAACCTCGAGTGGTGGATAGAGCAAGTCCAGCTTCC	413
Db	241	TAGAGGAACTGAGGAGAAAGGGAACCTCGAGTGGTGGATAGAGCAAGTCCAGCTTCC	300
QY	414	TGATCATGTGGCGAGGCCAGAAAGTGGGGAACCTCATCTATCAGTGGTTCACGGA	473
Db	301	TGATCATGTGGCGAGGCCAGAAAGTGGGGAACCTCATCTATCAGTGGTTCACGGA	360
QY	474	GTGGCCAGAACTCCGCTTTTACCTGTATGAATGAACTGACTAATGGGGAAGACACAGAGG	533
Db	361	GTGGCCAGAACTCCGCTTTTACCTGTATGAATGAACTGACTAATGGGGAAGACACAGAGG	420
QY	534	ATGAGGAGTTCACCGGCTGGATGAAGCCACTCTACTGGGGCTCTGCAAGGCTTACAGC	593
Db	421	ATGAGGAGTTCACCGGCTGGATGAAGCCACTCTACTGGGGCTCTGCAAGGCTTACAGC	480
QY	594	AGGAGCACAAAGCCGAGATCATCTGTGAGCGATGGCGAGGCGCTCAAGTTCTTTCAGC	653
Db	481	AGGAGCACAAAGCCGAGATCATCTGTGAGCGATGGCGAGGCGCTCAAGTTCTTTCAGC	540
QY	654	AGGAGCTGTCTCCCTTTTACCTTTTACCTCCCACTTTCAGGGCTTTTCAAAAGGAGACA	713
Db	541	AGGAGCTGTCTCCCTTTTACCTTTTACCTCCCACTTTCAGGGCTTTTCAAAAGGAGACA	600
QY	714	GACCCAGTGTCCCCAAAGACTGGATCTGTGATCCACGAGCTCAAAAGGACTCCAGTCC	773
Db	601	GACCCAGTGTCCCCAAAGACTGGATCTGTGATCCACGAGCTCAAAAGGACTCCAGTCC	660
QY	774	CTGAAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGCTCCCTTGGATAGG	833
Db	661	CTGAAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGCTCCCTTGGATAGG	720
QY	834	GTGAGGCTGAAGCACACCGGAGAAAATATGTGCTTTCTTCGCCCTTACCTCCTTTCCCAT	893


```

/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 102"
/note="Organ: salivary gland; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH MGC Library."

```

ORIGIN

[illegible]

```

Db      840  CACATGCGCTCGACACATGCTTGTGTTCTCCCTCCACCCCTTTAGTGTCT 999
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1022 GTTGCCTCCC 1031
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      900  GGTGCCCTCC 909
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 12
CO581180
LOCUS      936 bp      mRNA      linear      EST 20-JUL-2000
DEFINITION ILLUMIGEN MCQ_47660 Katze_MMDD Macaca mulatta cDNA clone
            IBUIW:18554 5' similar to Bases 718 to 934 highly similar to human
            MGC10540 (Hs:500165), mRNA sequence.
ACCESSION  CO581180
VERSION     CO581180
KEYWORDS   CO581180.1 GI:50412468
SOURCE     EST.
ORGANISM   Macaca mulatta (rhesus monkey)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE  1 (bases 1 to 936)
            Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
            Large-scale Rhesus Macaque cDNA Sequencing
            Unpublished (2003)
            Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.06.29. 724 Q20 bases. Library Preparation: Prof.
            Michael Katze Lab at University of Washington DNA Sequencing:
            Illumigen Biosciences Inc. For further information, see
            http://www.macaque.org
PCR Primers
FORWARD: CCCTCATAAAGGGAACAAA
BACKWARD: CACTATAGGCGAATGGGTA
Insert Length: 936 Std Error: 0.00
Plate: C1000407 row: F column: 12
Seq primer: CCCTCATAAAGGGAACAAA
POLYA=No.

FEATURES             Location/Qualifiers
     source            1..936
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBUIW:18554"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="Electromax DH10B"
                     /clone_lib="Katze MMDD"
                     /note="Organ: duodenum; Vector: pDONR 222; Site 1: BsrG I
                     Site 2: BsrG I; Created from CloneMiner cDNA Library
                     Construction kit (#18249-029)"

ORIGIN

Query Match      33.0%; Score 859; DB 7; Length 936;
Best Local Similarity 97.0%; Pred. No. 1.2e-177;
Matches 907; Conservative 0; Mismatches 25; Indels 3; Gaps 3

Qy      123  TGGCGATGAGTTTCGATGGCGCGTGGCAGTAGTCGCTCCACCCCTCTTACGTTACAC 182
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      2    TGGCGATGAGTTTCGATGGCGCGTGGCAGTAGTCGCTCCACCCCTCTTACGTTACAC 61
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Qy      183  CGAATGTGGACACTCGGCAGAACAGACTGGCGCGCTGGTGTGCTCGTGTGCTGTCTCTTCT 242
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      62   CGAATGTGGACACTCGGCAGAACAGACTGGCGCGCTGGTGTGCTCGTGTGCTGTCTCTCT 121
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Qy      243  GCOCGCTGCAAAACAGTCCAGCATGACGGTGATGGAAAGCTCAGAGAGCCCGCTCTTCA 302
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

Db	122	GCGGCTGCACAAACAGTCCAGATGACGGTGTGAGAGCTCAGGAGACCCGCTCTTCA	181
Qy	303	ACAACGTCAAGCTACAGCGAAAGCTTCCTGTGAGTGCATCCAGATTGATTAGAGGAC	362
Db	182	ACAACGTCAAGCTACAGCGAAAGCTTCCTGTGAGTGCATCCAGATTGATTAGAGGAC	241
Qy	363	TCAGGAGAAAGGACCTCGAGTGGTTCGATAAGAGCAAGTCCAGTTCCTGATCATGT	422
Db	242	TCAGGAGAAAGGACCTCGAGTGGTTCGATAAGAGCAAGTCTAGTTCCTGATCATGT	301
Qy	423	GCGGAGGCCAGAGAAAGTGGGAAACTCATCTATCAGTGGTTCACAGAGTGGCCACA	482
Db	302	GCGGAGGCCAGAGAAAGTGGGAAACTCATCTATCAGTGGTTCACAGAGTGGCCACA	361
Qy	483	ACAACCTCCGCTTTTACCCTGTATGAACACTAAATGGGGAAGACACAGAGGATGAGGAT	542
Db	362	ACAACCTCCGCTTTTACCCTGTATGAACACTAAATGGGGAAGACACAGAGGATGAGGAT	421
Qy	543	TCACGGGCTGATGAGGACACTTACTTCGGGGCTTCGAGGCCCTACAGCAGGAGCACAC	602
Db	422	TCATGCGGCTGATGAGGACACTTACTTCGGGGCTTCGAGGCCCTACAGCAGGAGCACAC	481
Qy	603	AGCGCGAGATCATCACTGTCAGCGATGGCGGAGGCTCAAGTTCCTTAGCAGGACCTG	662
Db	482	AGCGCGAGATCATCACTGTCAGCGATGGCGGAGGCTCAAGTTCCTTAGCAGGACCTG	541
Qy	663	TCCTCCCTTTACTTCTTACCTCCACCTTTCAGGGCTTTCAAAAGGAGACAGACCCAGTG	722
Db	542	TCCTCCCTTTACTTCTTACCTCCACCTTTCAGGGCTTTCAAAAGGAGACAGACCCAGTG	601
Qy	723	TCCTCCAAAGACTGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCTGAAGGCT	782
Db	602	TCCTCCAAAGACTGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCTGAAGGCT	661
Qy	783	GGGACTCGGGATGGTTCCTCACCCCATATGCTGTCCCTTGGATAGGGTGAAGGCTG	842
Db	662	GGGACTCGGGATGGTTCCTCACCCCATATGCTGTCCCTTGGATAGGGTGAAGGCTG	720
Qy	843	AAGCACCAGGAGAAAATATGTCTTCTTCGCGCTTACTCTCTTCCCATCTTAGACTG	902
Db	721	AAGCACCAGGAGAAAATATGTCTTCTTCGCGCTTACTCTCTTCCCATCTTAGACTG	779
Qy	903	TCCTTAGCCAGGGTCTGTAACCTGACACTTTATATGTTTCACACATGTAAGTACATA	962
Db	780	TCCTTAGCCAGGGTCTAT-AACTGTATCTTTATATGTTTCACACATGTAAGTACATA	838
Qy	963	CACATATGCGCTGCAGACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1022
Db	839	CACATATGCACTTGACAGGACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	898
Qy	1023	TTGCTCTCTCTCTCTCAGGCTGGTGGATCCTTCC 1057	
Db	899	TTGCTCTCTCTCTCAGGCTGGTGGATCCTTCTTC 933	
RESULT 13.			
BM919274			
LOCUS			
DEFINITION			
AGENCOURT_6715596 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748508			
5', mRNA sequence.			
ACCESSION			
BM919274.1 GI:19369653			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE			
1 (bases 1 to 1049)			
AUTHORS			
NIH-MGC http://mgi.nci.nih.gov/ .			
TITLE			
National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL			
Unpublished (1999)			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgabbs-r@mail.nih.gov			

Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM12776 row: 9 column: 05
High quality sequence stop: 650.
Location/Qualifiers
1. 1049
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5748508"
/lab_host="DH10B"
/clone_lib="NIH_MGC_120"
/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN

Query Match	32.8%;	Score	854.6;	DB	5;	Length	1049;
Best Local Similarity	97.8%;	Pred.	No. 1.2e-176;				
Matches	919;	Conservative	0;	Mismatches	15;	Indels	6;
Gaps	5;						
QY	110	TGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCTTC	169				
DB	1	TGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCTTC	60				
QY	170	TTTACGTTACACCGAATGTGGACACTCGGAGAGACAGCTGGCGCCCTGGTGTCTGCTG	229				
DB	61	TTTACGTTACACCGAATGTGGACACTCGGAGAGACAGCTGGCGCCCTGGTGTCTGCTG	120				
QY	230	GTCCGTGCTCTTCGCGCTGCACAAACAGTCCAGCATGACGCTGATGAAGCTCAGGAG	289				
DB	121	GTCCGTGCTCTTCGCGCTGCACAAACAGTCCAGCATGACGCTGATGAAGCTCAGGAG	180				
QY	290	AGCCGCTCTTCAACAAAGCTCAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGATT	349				
DB	181	AGCCGCTCTTCAACAAAGCTCAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGATT	240				
QY	350	GTATTAGAGAACTGAGGAGAAAGGAACTCTGAGTGGTGGATAAGAGCAAGTCCAGC	409				
DB	241	GTATTAGAGAACTGAGGAGAAAGGAACTCTGAGTGGTGGATAAGAGCAAGTCCAGC	300				
QY	410	TTCCCTGATCATGTGGCGGAGGCCAGAGAAATGGGGGAAACTCATCTATCAGTGGGTTTC	469				
DB	301	TTCCCTGATCATGTGGCGGAGGCCAGAGAAATGGGGGAAACTCATCTATCAGTGGGTTTC	360				
QY	470	AGGAGTGGCCAGAACTCCGCTTTTACCTGTATGAACACTGACTAAATGGGGAAGACACA	529				
DB	361	AGGAGTGGCCAGAACTCCGCTTTTACCTGTATGAACACTGACTAAATGGGGAAGACACA	420				
QY	530	GAGGATGAGGAGTTCACACGGGCTGGATGAAGCACTCTACTGCGGGGCTCTGCAGGGCCCTA	589				
DB	421	GAGGATGAGGAGTTCACACGGGCTGGATGAAGCACTCTACTGCGGGGCTCTGCAGGGCCCTA	480				
QY	590	CAGCAGGAGCAAGCGCCGAGATCATCTGTCCAGCGATGGCGGCGCTCAAGTCTTC	649				
DB	481	CAGCAGGAGCAAGCGCCGAGATCATCTGTCCAGCGATGGCGGCGCTCAAGTCTTC	540				
QY	650	TAGCAGGAGCACTCTCTCCCTTTTACTTCTTACCTCCACCTTTCCAGGGCTTTCAAAGGA	709				
DB	541	TAGCAGGAGCACTCTCTCCCTTTTACTTCTTACCTCCACCTTTCCAGGGCTTTCAAAGGA	600				
QY	710	GACAGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTCC	769				

REFERENCE 1 (bases 1 to 990)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMW12262 row: h column: 02
High quality sequence stop: 685.
FEATURES
source
1..990
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:550385"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH MGC 88"
/note="Organ: small intestine; Vector: pCMV-SPORT6; Site 1: Noti; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 32.4%; Score 843.8; DB 4; Length 990;
Best Local Similarity 98.5%; Pred. No. 2.7e-174;
Matches 871; Conservative 0; Mismatches 10; Indels 3; Gaps 2;
QY 153 ATCGCTTCCACCCCTCTTTAGCTTACACCGAATGTGGACACTCGGCAGAGAGCAGCTGG 212
DB 1 ATCGCTTCCACCCCTCTTTAGCTTACACCGAATGTGGACACTCGGCAGAGAGCAGCTGG 60
QY 213 CGGCTGTGTGCTCGCTGTCTCTTCTGCGGCTGCACAAAGTCCAGCATGACGG 272
DB 61 CGGCTGTGTGCTCGCTGTCTCTTCTGCGGCTGCACAAAGTCCAGCATGACGG 120
QY 273 TGATGGAAGCTCAGGAGAGCCGCTCTTCAACAAAGTCAAGCTACAGGAAAGCTTCCTG 332
DB 121 TGATGGAAGCTCAGGAGAGCCGCTCTTCAACAAAGTCAAGCTACAGGAAAGCTTCCTG 180
QY 333 TGGAGTCGATCCAGATTGTATTAGGAGACTGAGGAAGGAAGCACTCGAGTGGTTGG 392
DB 181 TGGAGTCGATCCAGATTGTATTAGGAGAACTGAGGAAGGAAGGAAGCACTCGAGTGGTTGG 240
QY 393 ATAAGAGCAAGTCCAGCTTCTGATCATGTGCGGAGGCGCAGAAAGTGGGGAAACTCA 452
DB 241 ATAAGAGCAAGTCCAGCTTCTGATCATGTGCGGAGGCGCAGAAAGTGGGGAAACTCA 300
QY 453 TCTATCAGTGGGTTTCCAGGAGTGGCCAGAACTCCGTCTTTTACCTGTATGAACCTGA 512
DB 301 TCTATCAGTGGGTTTCCAGGAGTGGCCAGAACTCCGTCTTTTACCTGTATGAACCTGA 360
QY 513 CTAATGGGGAAGACACAGAGGATGAGGATTCACGGGCTGGATGAGCCACTCTACTGC 572
DB 361 CTAATGGGGAAGACACAGAGGATGAGGATTCACGGGCTGGATGAGCCACTCTACTGC 420
QY 573 GGGCTCTGCAGGCCCTTACAGAGGAGCAAGCGGAGATCATCTGTCCAGGATGGCC 632
DB 421 GGGCTCTGCAGGCCCTTACAGAGGAGCAAGCGGAGATCATCTCCGTCCAGGATGGCC 480
QY 633 GAGGCGTCAAGTTCTTTAGCAGGAGCTGTCTCCCTTTTACTTTTACCTCCACCTTTC 692
DB 481 GAGGCGTCAAGTTCTTTAGCAGGAGCTGTCTCCCTTTTACTTTTACCTCCACCTTTC 540
QY 693 CAGGGCTTCAAAGGAGACAGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACC 752

Db 541 CAGGGCTTCAAAGGAGACAGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACC 600
QY 753 AGACTCAAAGGACTCCAGTCTCTGAAGGCTGGGACTGGGATGGGTTTCTCACACCCCA 812
Db 601 AGACTCAAAGGACTCCAGTCTCTGAAGGCTGGGACTGGGATGGGTTTCTCACACCCCA 660
QY 813 TATGTCTGCTCCCTTGGATAGGCTGGAAGCACCAGGGAGAAATATGTGCTTCTTC 872
Db 661 TATGTCTGCTCCCTTGGATAGGCTGGAAGCACCAGGGAGAAATATGTGCTTCTTC 720
QY 873 TCGCCCTACCTCTCTTCCCATCTTAGACTGTCTTTGAGCCAGGGTCTGTAAACCTGACAC 932
Db 721 TCGCCCTACCTCTCTTCCCATCTTAGACTGTCTTTGAGCCAGGGTCTGTAAACCTGACAC 780
QY 933 TTTATATGTGTTTCAACATGTAAAGTACATACATACATACATACATACATACATACATACAT 989
Db 781 TTTATATGTGTTTCAACATGTAAAGTACATACATACATACATACATACATACATACATACAT 840
QY 990 TGTCTCTCTCTCTCTCCGACCCCTTTAGCTGTGTTGCTCCCTT 1033
Db 841 GTNNTCTCTCTCTCTCCGACCCCTTTAGCTGTGTTGCTCCCTT 884

Search completed: June 29, 2005, 05:14:46
JOB time : 5615.03 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:57:12 ; Search time 8104.93 Seconds
(without alignments)
16865.305 Million cell updates/sec

Title: US-08-731-499-6
Perfect score: 2821
Sequence: 1 ATCTTAAGACGACAGCCTG.....AGGGTTCAAGACCAGCCTG 2821

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapept 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2821	100.0	2821	6	AR070329 Sequence
2	2821	100.0	2821	6	BD085730 Genes fro
3	2805.8	99.5	167078	9	HSJ139D8 Human DNA
4	1679.6	59.5	3008	9	HSGCAP13 Homo sapi
5	1635.2	58.0	2288	9	HSW805690 Homo sapi
6	956.6	33.9	2967	9	HSW807325 Homo sapi
7	374	13.3	407	11	G30061 human STS
8	337	11.9	1506	6	CQ491451 Sequence
9	337	11.9	1506	6	CQ497332 Sequence
10	278.6	9.9	33201	2	AL390735 Homo sapi
11	278	9.9	146530	9	AC023158 Homo sapi
12	277.8	9.8	74754	2	CR759924 Homo sapi
13	277.8	9.8	78827	9	CR788282 Human DNA
14	277.8	9.8	109067	9	CR388382 Human DNA
15	277.8	9.8	148281	2	CR759926 Homo sapi
16	277.8	9.8	187607	9	AC010329 Homo sapi
17	277.4	9.8	154957	9	AC019270 Homo sapi
18	277.2	9.8	56403	9	AJ418064 Homo sapi
19	277.2	9.8	164076	9	AC021059 Homo sapi

20	276.8	9.8	172981	9	AP000851 Homo sapi
21	276.8	9.8	185323	2	AC078970 Homo sapi
22	276.4	9.8	105580	9	Z82975 Human DNA s
23	276	9.8	138411	9	AL121958 Human DNA
24	275.8	9.8	71092	9	AL590732 Human DNA
25	275.8	9.8	83727	9	AL356120 Human DNA
26	275.8	9.8	114701	9	AC006927 Homo sapi
27	275.6	9.8	160396	9	AL592063 Human DNA
28	275.6	9.8	187350	9	AC020713 Homo sapi
29	275.2	9.8	173260	9	CNS078CU Human chr
30	275	9.7	38601	9	AC112646 Homo sapi
31	275	9.7	45120	9	AC090156 Homo sapi
32	275	9.7	114001	9	HSJ760G15 Human DNA
33	275	9.7	161943	9	AL356460 Homo sapi
34	275	9.7	165799	2	AC100813 Homo sapi
35	275	9.7	174373	9	AC132216 Homo sapi
36	274.6	9.7	111783	9	AL592449 Human DNA
37	274.6	9.7	144314	2	AP000835 Homo sapi
38	274.6	9.7	171732	9	AC022844 Homo sapi
39	274.6	9.7	176560	2	AC021852 Homo sapi
40	274.6	9.7	180019	9	AC013734 Homo sapi
41	274.6	9.7	180547	9	AL139331 Human DNA
42	274.6	9.7	186152	2	AC130358 Homo sapi
43	274.6	9.7	192070	9	AC090616 Homo sapi
44	274.2	9.7	107685	9	EX679664 Human DNA
45	274.2	9.7	129819	9	AC104049 Homo sapi

ALIGNMENTS

RESULT 1
AR070329
LOCUS AR070329 2821 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 6 from patent US 5892010.
ACCESSION AR070329
VERSION AR070329.1 GI:7221217
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2821)
AUTHORS Gray,J., Collins,C., Hwang,S.-i., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20Q13 amplicon and their uses
JOURNAL Patent: US 5892010-A 6 06-APR-1999;
FEATURES
source Location/Qualifiers
1..2821
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN	Query Match	100.0%;	Score 2821;	DB 6;	Length 2821;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2821;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	ATCCTTAAGACGACAGCCTTGGGAAGCTGCTGGGGAAGTGGTCTGAGGATGTGGGT	60		
Db	1	ATCCTTAAGACGACAGCCTTGGGAAGCTGCTGGGGAAGTGGTCTGAGGATGTGGGT	60		
Qy	61	CACCTGGGTGAAGGTGGAGCTTTTCAGGGTCTCCCGTCAATGACGTAGTATTTCTTTGGC	120		
Db	61	CACCTGGGTGAAGGTGGAGCTTTTCAGGGTCTCCCGTCAATGACGTAGTATTTCTTTGGC	120		
Qy	121	AGGGAATTACCACTGAAGAGCTCCCGCGAGAGCTACAACTGACGAGCCAGC	180		
Db	121	AGGGAATTACCACTGAAGAGCTCCCGCGAGAGCTACAACTGACGAGCCAGC	180		
Qy	181	TGCTCACACCCGAGGAGGTCTGGGACAGATCTTCTCTCTGTTGGATGAGATGAGATG	240		
Db	181	TGCTCACACCCGAGGAGGTCTGGGACAGATCTTCTCTCTGTTGGATGAGATGAGATG	240		
Qy	241	GTAAGAGGGGCGAGAGATGGGGAGAGTGTGTCTCACTCTGTCATCATCGGCCACTTTCTGGCC	300		

|||||
241 GTAAAGGGCAGAGATGGGAGAGTGTGTCCACTCTGCAATCATCGCCACTTTCTGGCC 300
QY
301 GCACGTCTCTGGGCAAGGCCCTCACCCTTCCAAACCTCGGGTCCCTCATCTGTGTGAAGAGC 360
DB
301 GCACGTCTCTGGGCAAGGCCCTCACCCTTCCAAACCTCGGGTCCCTCATCTGTGTGAAGAGC 360
QY
361 TGTGAGAAGATGTCTCACTGAACCTAAACAAGGGACTCATAGACACGTGTTTGTAGGAGTGCAC 420
DB
361 TGTGAGAAGATGTCTCATGAACTTAACAAGGGACTCATAGACACGTGTTTGTAGGAGTGCAC 420
QY
421 TAAAGTCTCTACAGGAGTTGCTGATGGAGGCCAGGCACGCAGAAATAGAAAGATAGGAAAC 480
DB
421 TAAAGTCTCTACAGGAGTTGCTGATGGAGGCCAGGCACGCAGAAATAGAAAGATAGGAAAC 480
QY
481 TTTGAGTCTAGCCAGGAGTGATATATTGAGGCTTCTCGTCTAGTCTCAATTTCTCATC 540
DB
481 TTTGAGTCTAGCCAGGAGTGATATATTGAGGCTTCTCGTCTAGTCTCAATTTCTCATC 540
QY
541 TGGAAATGGGGATAATAATAGTGTGTGAGAGGAATGAATAGGATAATGTGTTTAAGAGC 600
DB
541 TGGAAATGGGGATAATAATAGTGTGTGAGAGGAATGAATAGGATAATGTGTTTAAGAGC 600
QY
601 AGGCATAGGGTAGACCTCCATTACAGGCTGCTTTGGGCTTTTCCCTCTGTAGCCCAAGGCC 660
DB
601 AGGCATAGGGTAGACCTCCATTACAGGCTGCTTTGGGCTTTTCCCTCTGTAGCCCAAGGCC 660
QY
661 AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGAAATACACACTTGAAGCCCTCAGCTC 720
DB
661 AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGAAATACACACTTGAAGCCCTCAGCTC 720
QY
721 TCTCAGCTCCACCAGCATTTCCGTGTACCATCGCAAAAGTAAATCTCAATTCATCA 780
DB
721 TCTCAGCTCCACCAGCATTTCCGTGTACCATCGCAAAAGTAAATCTCAATTCATCA 780
QY
781 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTCTGGAGATTTCAGACACCTGCAAGGCC 840
DB
781 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTCTGGAGATTTCAGACACCTGCAAGGCC 840
QY
841 GCAAGCTTGAGGTCTGAGACTCAGAGATAGTTGGCATTAAGAGCTGAGCGCATCTGGG 900
DB
841 GCAAGCTTGAGGTCTGAGACTCAGAGATAGTTGGCATTAAGAGCTGAGCGCATCTGGG 900
QY
901 AGCAGGTCTCTCTCTGCCCCCAGACCCAGGTCTCCCTTCTTCTACATGACCACTC 960
DB
901 AGCAGGTCTCTCTCTGCCCCCAGACCCAGGTCTCCCTTCTTCTACATGACCACTC 960
QY
961 TCCTCCCCCTTGTCTCAGGCCAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCCCGTCGGGA 1020
DB
961 TCCTCCCCCTTGTCTCAGGCCAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCCCGTCGGGA 1020
QY
1021 CAAAGTGGGTGATGAAGATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCA 1080
DB
1021 CAAAGTGGGTGATGAAGATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCA 1080
QY
1081 GAGACGGAAAGTGCCATGTTCTGAGGAGTCTGGGGCCCTTCCAGACTCCAGGCTCACCC 1140
DB
1081 GAGACGGAAAGTGCCATGTTCTGAGGAGTCTGGGGCCCTTCCAGACTCCAGGCTCACCC 1140
QY
1141 CAGGTTTCCAGGGTAGTAGGAGGGTCCCCTGCTCAGCTGTCTCATGCCCCACTCTTCCCC 1200
DB
1141 CAGGTTTCCAGGGTAGTAGGAGGGTCCCCTGCTCAGCTGTCTCATGCCCCACTCTTCCCC 1200
QY
1201 TGGTGTTCAGTCTCTGGCACCCTTGTGAGCGGCTGAGTGGGGATGGGAGGGCTGCTG 1260
DB
1201 TGGTGTTCAGTCTCTGGCACCCTTGTGAGCGGCTGAGTGGGGATGGGAGGGCTGCTG 1260
QY
1261 GGTTTGAAGTGCCCAACAGGGCATAGTCCATTTTCGAGAGTCCCTCGGATGGTGAAGGG 1320
DB
1261 GGTTTGAAGTGCCCAACAGGGCATAGTCCATTTTCGAGAGTCCCTCGGATGGTGAAGGG 1320
QY
1321 AATTCAAGTTACTTTTCTGTTTTCAGCCGCTCTCGGAGGACTGTGCTGCTTGGGTGGGT 1380
|||||

DB
1321 AATTCAAGTTACTTTTCTGTTTTCAGCCGCTCTCTGGAGGACTGTGCTTGGCTGGGTGGTT 1380
QY
1381 GTGGGCTCCCAACAGTTTCTCGGTGTTCTCAGTTTGGAGCAGAGCCAACTCAGAGGGTGA 1440
DB
1381 GTGGGCTCCCAACAGTTTCTCGGTGTTCTCAGTTTGGAGCAGAGCCAACTCAGAGGGTGA 1440
QY
1441 GGGTCCCAACAGACCAAAATCAGAAATAGAAACACAAAGACTGTTAGAGAGCGAGGGTGGGA 1500
DB
1441 GGGTCCCAACAGACCAAAATCAGAAATAGAAACACAAAGACTGTTAGAGAGCGAGGGTGGGA 1500
QY
1501 GGGTGTTCAGACTGAAGAAAAGGAGAGTTGCCGGGCAACGGTGGCTCAGCCTGTAATC 1560
DB
1501 GGGTGTTCAGACTGAAGAAAAGGAGAGTTGCCGGGCAACGGTGGCTCAGCCTGTAATC 1560
QY
1561 CCAGCATTTTGGGAGCGCGAGCGGCGGAGATACAGAGGTGAGGAGATCAGAGCAACCATCTG 1620
DB
1561 CCAGCATTTTGGGAGCGCGAGCGGCGGAGATACAGAGGTGAGGAGATCAGAGCAACCATCTG 1620
QY
1621 GCTAACACGGGTGAAAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGGTGAGGTG 1680
DB
1621 GCTAACACGGGTGAAAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGGTGAGGTG 1680
QY
1681 GCGGCGGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGAAATGGCGTGAACCCCA 1740
DB
1681 GCGGCGGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGAAATGGCGTGAACCCCA 1740
QY
1741 GGGGCGGAGCCTACAGTGAGCCGAGATTGCGCACCTGCATCTCCAGCCTGGACGACAGTG 1800
DB
1741 GGGGCGGAGCCTACAGTGAGCCGAGATTGCGCACCTGCATCTCCAGCCTGGACGACAGTG 1800
QY
1801 AGACTCCGCTCTCAAAAAAAGAAAGAAAGAAAGGACGAGTGTGGGGGGCAG 1860
DB
1801 AGACTCCGCTCTCAAAAAAAGAAAGAAAGAAAGGACGAGTGTGGGGGGCAG 1860
QY
1861 GGGGAGCAATTAATTTCTATAACTTCCTGGGATGCTGAGGGGCGTTCATGGGAGGACCCCTG 1920
DB
1861 GGGGAGCAATTAATTTCTATAACTTCCTGGGATGCTGAGGGGCGTTCATGGGAGGACCCCTG 1920
QY
1921 GCCTCTCTCTCCCAAGCATCTCCAGGTGTGTCAACAGGAAATGGCAGCAATA 1980
DB
1921 GCCTCTCTCTCCCAAGCATCTCCAGGTGTGTCAACAGGAAATGGCAGCAATA 1980
QY
1981 CGCTCAGGCTGTGTCTTTCTGCTTTGAAAGGCTCAGCTGTACTTAAAGGACTGT 2040
DB
1981 CGCTCAGGCTGTGTCTTTCTGCTTTTGAAGGCTCAGCTGTACTTAAAGGACTGT 2040
QY
2041 CAGCTCTGCTGGGTGCTGTGGGACCCCTGCTGCAACCCACCACTCCCCCAACAA 2100
DB
2041 CAGCTCTGCTGGGTGCTGTGGGACCCCTGCTGCAACCCCACTCCCCCAACAA 2100
QY
2101 TCCTCTCTTTCCATCCATATCCCGAGTATGGACCTTCCAACTCCAGCCATAGCTG 2160
DB
2101 TCCTCTCTTTCCATCCATATCCCGAGTATGGACCTTCCAACTCCAGCCATAGCTG 2160
QY
2161 AATGTTTCTTTTAAAGGATGGAGAAAATCTTCTGTCTCTCTGCAAGAAATGGGGGAC 2220
DB
2161 AATGTTTCTTTTAAAGGATGGAGAAAATCTTCTGTCTCTCTGCAAGAAATGGGGGAC 2220
QY
2221 TGTGACTGGGATGTGGGCTGGGCTTGGCTTCTTAACTGTGTGACCAAGACAGCA 2280
DB
2221 TGTGACTGGGATGTGGGCTGGGCTTGGCTTCTTAACTGTGTGACCAAGACAGCA 2280
QY
2281 CTTCTCTCTCCCTTAACTTTGTTTGTGTCAGCAGCAGTGGAGGTCGACTAGGCGA 2340
DB
2281 CTTCTCTCTCCCTTAACTTTGTTTGTGTCAGCAGCAGTGGAGGTCGACTAGGCGA 2340
QY
2341 ACAGTTTGGATTAATTTGTTTGTAGATGTGGAATTAATTTTGTATATAAATCTTAT 2400
DB
2341 ACAGTTTGGATTAATTTGTTTGTAGATGTGGAATTAATTTTGTATATAAATCTTAT 2400
QY
2401 GTGTAAACCCCAATATAGAACTAGATTAAGAGGAGTCTCTCTGTTGAAGGGGAGCTG 2460
DB
2401 GTGTAAACCCCAATATAGAACTAGATTAAGAGGAGTCTCTCTGTTGAAGGGGAGCTG 2460


```

QY 2461 AGTACCTCTGGAACCTGGAGGCACTCTGAAAAAAGCAAACTGAAAAACCAAGTGCCTGGG 2520
DB 2461 AGTACCTCTGGAACCTGGAGGCACTCTGAAAAAAGCAAACTGAAAAACCAAGTGCCTGGG 2520
QY 2521 TCACCTGTTACTCTATAGACAGTTTAAAGTCAGACCTGGAAAAACATTTGCTTTACCTT 2580
DB 2521 TCACCTGTTACTCTATAGACAGTTTAAAGTCAGACCTGGAAAAACATTTGCTTTACCTT 2580
QY 2581 GAATAGATAGGTTTTTATGCTTGGTATATAAGAAATAAAACTAACTATTAAACCTTGAGAC 2640
DB 2581 GAATAGATAGGTTTTTATGCTTGGTATATAAGAAATAAAACTAACTATTAAACCTTGAGAC 2640
QY 2641 TTACAGGTGCTTATTTATCATATGATAGTCAATATAAAATTTCTTTAGACATCAATTTTA 2700
DB 2641 TTACAGGTGCTTATTTATCATATGATAGTCAATATAAAATTTCTTTAGACATCAATTTTA 2700
QY 2701 GGTAAAAATANTGATTAGAAAAATATTTGGCCAGGTGCGAGCTCACACCTGCAATCC 2760
DB 2701 GGTAAAAATANTGATTAGAAAAATATTTGGCCAGGTGCGAGCTCACACCTGCAATCC 2760
QY 2761 CAGGACTTTGGAGGCCGAGGGCGGTGGATCACTGAGGTCAAGGCTTCAAGACCAAGCT 2820
DB 2761 CAGGACTTTGGAGGCCGAGGGCGGTGGATCACTGAGGTCAAGGCTTCAAGACCAAGCT 2820
QY 2821 G 2821
DB 2821 G 2821

RESULT 2
BD085730
LOCUS BD085730 2821 bp DNA linear PAT 27-AUG-2002
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085730
VERSION BD085730.1 GI:22631340
KEYWORDS JP 2001524802-A/6.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 2821)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and
Romens,J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 6 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
COMMENT OS Artificial Sequence
PN JP 2001524802-A/6
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
PI 17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11, C12Q1/68, A61K48/00
CC Description of Artificial Sequence:CCAP
FH Key Location/Qualifiers
FT source 1..2821
FT /organism='Artificial Sequence'.
FEATURES
source
1..2821
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'

ORIGIN
Query Match 100.0%; Score 2821; DB 6; Length 2821;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2821; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCCTAAGACGCACAGCCTGGGAAGCCAGACA CTGGGGAAGTGGTGTCTGAGGGAATGTGGGT 60

```

```

DB 1 ATCCTAAGACGCACAGCCTGGGAAGCCAGACA CTGGGGAAGTGGTGTCTGAGGGAATGTGGGT 60
QY 61 CACTGGGTGGAAGGTGGAGCTTTTCAAGGGTCTCCCGTCAATGACAGCTGAGTTTCTTTGGC 120
DB 61 CACTGGGTGGAAGGTGGAGCTTTTCAAGGGTCTCCCGTCAATGACAGCTGAGTTTCTTTGGC 120
QY 121 AGGGAATTTACAGCTGAAGAAAGCCTGCGGCGAGAGCTACAAACTGAGCAAGGCCAGC 180
DB 121 AGGGAATTTACAGCTGAAGAAAGCCTGCGGCGAGAGCTACAAACTGAGCAAGGCCAGC 180
QY 181 TGCTCACACCCGAGGAGGTGCTGGACAGATCTTCTCTCTGCTGGATGAGAAATGGAATG 240
DB 181 TGCTCACACCCGAGGAGGTGCTGGACAGATCTTCTCTCTGCTGGATGAGAAATGGAATG 240
QY 241 GTAAGAGGGGACAGATGAGGAGAGTGTGTCTCACTCTGCAATCATCGCCACCTTTCTGGCC 300
DB 241 GTAAGAGGGGACAGATGAGGAGAGTGTGTCTCACTCTGCAATCATCGCCACCTTTCTGGCC 300
QY 301 GCACGTCTCTGGGCAAGGCCCTTCACTTCCAACTTGGGGTCTCTCATCTGTGAGAAAGCC 360
DB 301 GCACGTCTCTGGGCAAGGCCCTTCACTTCCAACTTGGGGTCTCTCATCTGTGAGAAAGCC 360
QY 361 TGTGAGAGAGATGTCATGAACCTAAACAAAGGAGCTCATGAGCAGCTGTTTGTAGGAGTGAC 420
DB 361 TGTGAGAGAGATGTCATGAACCTAAACAAAGGAGCTCATGAGCAGCTGTTTGTAGGAGTGAC 420
QY 421 TAAAGTCTCTACAGGAGTTGCTGATGGAGGCCAGGACGACGACAGATAGAAAGAAATAGGAAC 480
DB 421 TAAAGTCTCTACAGGAGTTGCTGATGGAGGCCAGGACGACGACAGATAGAAAGAAATAGGAAC 480
QY 481 TTTGAGTCAAGGACGAGGAGTGATATATTGAGGCTTCTGCTCTAGTCTCAATTTCTCTCATC 540
DB 481 TTTGAGTCAAGGACGAGGAGTGATATATTGAGGCTTCTGCTCTAGTCTCAATTTCTCTCATC 540
QY 541 TGGAAATGGGGATAATATAGTGGTTCAGAGGATGAATAGGATAATGTTTAAAGAGC 600
DB 541 TGGAAATGGGGATAATATAGTGGTTCAGAGGATGAATAGGATAATGTTTAAAGAGC 600
QY 601 AGGCATAGGGTAGACCTCCATTCAGGCTGCTTGGGCTTTCTCCCTGTAGGCCAAAGCCC 660
DB 601 AGGCATAGGGTAGACCTCCATTCAGGCTGCTTGGGCTTTCTCCCTGTAGGCCAAAGCCC 660
QY 661 AGCCTCAGGGCTATGTTGGGAGAGAGCTGGCTTGGAAATACACACTTGGCCCTCCAGCTC 720
DB 661 AGCCTCAGGGCTATGTTGGGAGAGAGCTGGCTTGGAAATACACACTTGGCCCTCCAGCTC 720
QY 721 TCTCAGCTCCACCCAGCATTTCCGTGTACCATGCGCAAAAGTAAACTTCAATTCATCA 780
DB 721 TCTCAGCTCCACCCAGCATTTCCGTGTACCATGCGCAAAAGTAAACTTCAATTCATCA 780
QY 781 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCCTGGAGATTTCAGACACCTGACAAGCC 840
DB 781 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCCTGGAGATTTCAGACACCTGACAAGCC 840
QY 841 GCAAGCTTTGAGGTCTGAGACTGACAGATAGTTGGCATAGACGTGAGCGCATCTCTGGG 900
DB 841 GCAAGCTTTGAGGTCTGAGACTGACAGATAGTTGGCATAGACGTGAGCGCATCTCTGGG 900
QY 901 AGCGAGGTCTCTCTCTCCCGCCAGACCCAGGCTCTCCCTCTTCTACATGACCACTC 960
DB 901 AGCGAGGTCTCTCTCTCCCGCCAGACCCAGGCTCTCCCTCTTCTACATGACCACTC 960
QY 961 TCCTCCCTCTGCTCAGGCCAGCTCTCTCTGAAACAGATTGTTGAAGGTGCCCGTCGGGA 1020
DB 961 TCCTCCCTCTGCTCAGGCCAGCTCTCTCTGAAACAGATTGTTGAAGGTGCCCGTCGGGA 1020
QY 1021 CAAGTGGGTGATGAAGATGCTGACATGGAATGAAATCCCAGCAGCTGGCTCGCTCAGCA 1080
DB 1021 CAAGTGGGTGATGAAGATGCTGACATGGAATGAAATCCCAGCAGCTGGCTCGCTCAGCA 1080
QY 1081 GAGACGGAAAAGTGCATGTTCTGAGGAGTCTGGGGGCCCTCCACGACTCCAGGCTCACC 1140
DB 1081 GAGACGGAAAAGTGCATGTTCTGAGGAGTCTGGGGGCCCTCCACGACTCCAGGCTCACC 1140

```

1141 CAGGTTTCCAGGGTAGTAGAGGGTCCCTCGCTCAGCCTGCTCATGCCCCACTCTTCCCC 1200
1141 CAGGTTTCCAGGGTAGTAGAGGGTCCCTCGCTCAGCCTGCTCATGCCCCACTCTTCCCC 1200
1201 TGGTGTGACTTCTCTGGCACCCCTGTGTGAGGGCTGAGTGGGGATGGGGAAGGGCTGCTG 1260
1201 TGGTGTGACTTCTCTGGCACCCCTGTGTGAGGGCTGAGTGGGGATGGGGAAGGGCTGCTG 1260
1261 GGTTTGAAGTGGCCAAAGGGCATAGTCATTTTGGAGGAGTCCCTGGGATGGTGAAGGG 1320
1261 GGTTTGAAGTGGCCAAAGGGCATAGTCATTTTGGAGGAGTCCCTGGGATGGTGAAGGG 1320
1321 AATTCAAGTACTTTTCTGTTTCAGCCGCTCTCTGGGAGGACTGTGCTTGGCTGGGTGGTT 1380
1321 AATTCAAGTACTTTTCTGTTTCAGCCGCTCTCTGGGAGGACTGTGCTTGGCTGGGTGGTT 1380
1381 GTGGGCTCCACAGATTTTCTGGGTGTTCTCAGTTTGGAAAGAGAGCCAACTGAGGGGTGA 1440
1381 GTGGGCTCCACAGATTTTCTGGGTGTTCTCAGTTTGGAAAGAGAGCCAACTGAGGGGTGA 1440
1441 GGGTCCCAACAGCCAAATCAGAAATGAGAAACAAAGACTGGTAGGAGGCGAGGGTGGGA 1500
1441 GGGTCCCAACAGCCAAATCAGAAATGAGAAACAAAGACTGGTAGGAGGCGAGGGTGGGA 1500
1501 GGGTGTTCAGACTGGAAGAAAAGGCGAGGAGTTGCGGGCACGCTGGCTCACGCTGTGAATC 1560
1501 GGGTGTTCAGACTGGAAGAAAAGGCGAGGAGTTGCGGGCACGCTGGCTCACGCTGTGAATC 1560
1561 CCAGCACTTTGGAGGCGGAGGCGGCGAGATCAGAGGTCAGGAGATCGAGACCATCCTG 1620
1561 CCAGCACTTTGGAGGCGGAGGCGGCGAGATCAGAGGTCAGGAGATCGAGACCATCCTG 1620
1621 GCTAACACGGGTGAAACCCCGTCTCTACTTAAATAACAAATAACAAATAACAAATAAC 1680
1621 GCTAACACGGGTGAAACCCCGTCTCTACTTAAATAACAAATAACAAATAACAAATAAC 1680
1681 GCGGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGCAAGAGATGCGCTGGAACCCCA 1740
1681 GCGGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGCAAGAGATGCGCTGGAACCCCA 1740
1741 GGGGGCCGAGCTACAGTGAGCCGAGATTGCGCACTGCACTCCAGCCTGGACGACAGTG 1800
1741 GGGGGCCGAGCTACAGTGAGCCGAGATTGCGCACTGCACTCCAGCCTGGACGACAGTG 1800
1801 AGACTCCGTCTCAAAAAAAGAAAGAAAGAAAGGAGGAGTGTGGGGGGCAG 1860
1801 AGACTCCGTCTCAAAAAAAGAAAGAAAGAAAGGAGGAGTGTGGGGGGCAG 1860
1861 GGGGAGCAATAATTTCTATACTTCCGGGATGCTGAGGGGGTTCATGGGAGGACCCCTG 1920
1861 GGGGAGCAATAATTTCTATACTTCCGGGATGCTGAGGGGGTTCATGGGAGGACCCCTG 1920
1921 GCCTCTCTCTCCCAAGGATCTCTCACAGTGGTGTCAACAGGAAAAATGCGAGCAATA 1980
1921 GCCTCTCTCTCCCAAGGATCTCTCACAGTGGTGTCAACAGGAAAAATGCGAGCAATA 1980
1981 CGCTGAGGCTGTGGTCTTTCTGCTTTGAAAGGTCAGCTGTACTTAAAGGAGTGTGT 2040
1981 CGCTGAGGCTGTGGTCTTTCTGCTTTGAAAGGTCAGCTGTACTTAAAGGAGTGTGT 2040
2041 CAGCTCTCCCTGGGTGCTGCTGGGACCCCTGCTGCAACCCCAACTCTCCCCCAACA 2100
2041 CAGCTCTCCCTGGGTGCTGCTGGGACCCCTGCTGCAACCCCAACTCTCCCCCAACA 2100
2101 TCCTCTCTTTCATCATATCCCCAGTATGACCTTCCCAACTCCAGGCATAAGCTG 2160
2101 TCCTCTCTTTCATCATATCCCCAGTATGACCTTCCCAACTCCAGGCATAAGCTG 2160
2161 AATGTTTCTTTAAGATGAGGAAACTTCTGCTCTCTGCGAGAAATTTGGGGAC 2220
2161 AATGTTTCTTTAAGATGAGGAAACTTCTGCTCTCTGCGAGAAATTTGGGGAC 2220

2221 TGTGACTGGGATTTGTGGCTGGGCTTAACTGCTGTGTGAACCAAGACAGCCA 2280
2221 TGTGACTGGGATTTGTGGCTGGGCTTAACTGCTGTGTGAACCAAGACAGCCA 2280
2281 CTTCTCTCTCCCTAACTTGTGTCTTGGCAGCACAGTGAAGGCTGGAGCTAGGCGA 2340
2281 CTTCTCTCTCCCTAACTTGTGTCTTGGCAGCACAGTGAAGGCTGGAGCTAGGCGA 2340
2341 ACAGTTTGGATTTAGTGTGTTTGTAGATGGAATTTTGTATATAAATCTTAT 2400
2341 ACAGTTTGGATTTAGTGTGTTTGTAGATGGAATTTTGTATATAAATCTTAT 2400
2401 GTGTAAACCCCAATATAGAACTAGATTAAAGAGGAGTCTCTCTGTTGAAAGGGGAGCTG 2460
2401 GTGTAAACCCCAATATAGAACTAGATTAAAGAGGAGTCTCTCTGTTGAAAGGGGAGCTG 2460
2461 AGTACCTCTCTGAACTGAGGACCTCTGTAAGAAAGCAAACTGAAACCCAGTGGCTGGG 2520
2461 AGTACCTCTCTGAACTGAGGACCTCTGTAAGAAAGCAAACTGAAACCCAGTGGCTGGG 2520
2521 TCAGTGTACTCTTATAGACAGTTTAAAGTGAGACCTGGAAACATTTGCTTTACCTT 2580
2521 TCAGTGTACTCTTATAGACAGTTTAAAGTGAGACCTGGAAACATTTGCTTTACCTT 2580
2581 GAATAGATAGTGTTTTATGTTGTTATATAAGAAATAAACTAAACCTTAAACCTTGAGAC 2640
2581 GAATAGATAGTGTTTTATGTTGTTATATAAGAAATAAACTAAACCTTAAACCTTGAGAC 2640
2641 TTTACAGTGTGTTTATTTATGATGATCATATAAAATTTTCTTTAGACATCAATTTTA 2700
2641 TTTACAGTGTGTTTATTTATGATGATCATATAAAATTTTCTTTAGACATCAATTTTA 2700
2701 GGTAAAAAATAATTTAGTAAATAATTTGGCCAGGTCAGACGCTCACCTGCAATCC 2760
2701 GGTAAAAAATAATTTAGTAAATAATTTGGCCAGGTCAGACGCTCACCTGCAATCC 2760
2761 CAGGACTTTGGGAGCGCGGCTGGATCATCCCTGAGGTCAGGGGTTCAAGACAGAGCT 2820
2761 CAGGACTTTGGGAGCGCGGCTGGATCATCCCTGAGGTCAGGGGTTCAAGACAGAGCT 2820
2821 G 2821
2821 G 2821

RESULT 3
HSDJ139D8
LOCUS
DEFINITION
HSDJ139D8 167078 bp DNA linear PRI 05-JUN-2003
Human DNA sequence from clone RPI-139D8 on chromosome 6p12.1-21.1
Contains genes for GUCA1A (guanylate cyclase activator 1A) and
GUCA1B (guanylate cyclase activator 1B) both from retina, the 3'
end of the fRep-132 gene for zinc finger transcription regulating
protein, the MRP510 gene for mitochondrial ribosomal protein S10,
the 5' end of the gene for a novel protein, the gene for a novel
protein and two CpG Islands, complete sequence.

ACCESSION
AL096814
VERSION
AL096814.26 GI:7159757
KEYWORDS
HTG; CpG Island; guanylate cyclase activator; GUCA1A; GUCA1B;
MRP510; fRep-132.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 6, 2000 this sequence version replaced gi:7018706.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6>

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

RP1-139D8 is from the library RPC1-1 constructed by the group of

Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2

This sequence is the entire insert of clone RP1-139D8 The true

right end of clone RP1-321B9 is at 5591 in this sequence.

FEATURES

source

1. 167078

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosomes="6"

/map="p12.1-21.1"

/clone="RP1-139D8"

/clone_lib="RPC1-1"

repeat_region

/note="3.8 copies 4 mer AGTT 23% conserved"

repeat_region

/note="2.0 copies 7 mer GTTAGTT 28% conserved"

repeat_region

/note="L3 repeat: matches 1197. 1345 of consensus"

repeat_region

/note="2.2 copies 5 mer CTCTG 22% conserved"

repeat_region

/note="2.5 copies 8 mer TTTGAAGA 40% conserved"

repeat_region

/note="2.8 copies 4 mer TCAT 22% conserved"

repeat_region

/note="L2 repeat: matches 3210. 3313 of consensus"

repeat_region

/note="L1MED repeat: matches 2400. 2696 of consensus"

repeat_region

/note="2.2 copies 5 mer TCATT 22% conserved"

repeat_region

/note="12.2 copies 4 mer GATA 91% conserved"

repeat_region

/note="2.5 copies 4 mer AGCT 20% conserved"

repeat_region

/note="7.0 copies 2 mer AT 28% conserved"

repeat_region

/note="14.0 copies 4 mer TAGA 53% conserved"

repeat_region

/note="AluJb repeat: matches 195. 288 of consensus"

repeat_region

/note="3.7 copies 3 mer CCT 22% conserved"

repeat_region

/note="7.8 copies 5 mer TTTTG 44% conserved"

repeat_region

/note="9.2 copies 4 mer TTTG 40% conserved"

repeat_region

/note="AluSx repeat: matches 5. 312 of consensus"

repeat_region

/note="1.9 copies 135 mer

CAACCATTCAGCTCCCAAGTGCTGCATTACGAGCATGACCGTGGCCAC

C 373% conserved"

repeat_region

/note="AluJb repeat: matches 1. 195 of consensus"

unsure

repeat_region

/note="L1MED repeat: matches 1010. 1074 of consensus"

repeat_region

/note="AluSx repeat: matches 1. 312 of consensus"

repeat_region

/note="2.2 copies 6 mer TGGAGG 26% conserved"

repeat_region

/note="36.0 copies 1 mer A 36% conserved"

repeat_region

/note="6.5 copies 6 mer AAAAAC 60% conserved"

repeat_region

/note="L1MED repeat: matches 318. 1010 of consensus"

repeat_region

/note="2.4 copies 7 mer TTTTGC 25% conserved"

repeat_region

/note="2.8 copies 4 mer TTTA 22% conserved"

repeat_region

/note="2.2 copies 5 mer GTGGA 22% conserved"

repeat_region

/note="2.0 copies 6 mer ATTCT 24% conserved"

repeat_region

/note="3.2 copies 5 mer CCTCT 32% conserved"

repeat_region

/note="2.0 copies 6 mer CCTGAA 24% conserved"

repeat_region

/note="2.4 copies 8 mer TCTCAGCT 29% conserved"

repeat_region

/note="2.4 copies 5 mer CTTC 24% conserved"

repeat_region

/note="2.0 copies 6 mer TTTTCC 24% conserved"

repeat_region

/note="MIR repeat: matches 8. 261 of consensus"

repeat_region

/note="3.1 copies 7 mer CTGTGCC 26% conserved"

repeat_region

/note="2.3 copies 6 mer TGGGSC 28% conserved"

repeat_region

/note="L1MC1 repeat: matches 6287. 6329 of consensus"

repeat_region

/note="L2 repeat: matches 3016. 3054 of consensus"

repeat_region

/note="2.2 copies 5 mer CCAGC 22% conserved"

repeat_region

/note="3.8 copies 5 mer CCCAC 29% conserved"

repeat_region

/note="2.0 copies 5 mer ATTAT 20% conserved"

repeat_region

/note="AluSg repeat: matches 1. 299 of consensus"

repeat_region

/note="19.0 copies 1 mer A 38% conserved"

repeat_region

/note="2.8 copies 13 mer CTACCCCTACTTC 56% conserved"

repeat_region

/note="2.2 copies 5 mer TCTCT 22% conserved"

repeat_region

/note="MIR repeat: matches 33. 73 of consensus"

repeat_region

/note="2.0 copies 8 mer ATATCTTG 32% conserved"

repeat_region

/note="MIR3 repeat: matches 85. 170 of consensus"

repeat_region

/note="2.3 copies 7 mer CAGGGCA 23% conserved"

repeat_region	6462. .6478	
repeat_region	/notes="4.2 copies 4 mer GGGT 27% conserved"	
repeat_region	6581. .6592	
repeat_region	/notes="2.0 copies 6 mer GGAGAG 24% conserved"	
repeat_region	7645. .7655	
repeat_region	/notes="5.5 copies 2 mer CT 22% conserved"	
repeat_region	7970. .7980	
repeat_region	/notes="2.2 copies 5 mer AACCA 22% conserved"	
repeat_region	8029. .8041	
repeat_region	/notes="2.6 copies 5 mer CAAAA 26% conserved"	
repeat_region	8047. .8057	
repeat_region	/note="2.2 copies 5 mer ACCCC 22% conserved"	
repeat_region	complement(8111..8281)	
repeat_region	/notes="MIR repeat: matches 20. .192 of consensus"	
repeat_region	8146. .8156	
repeat_region	/notes="3.7 copies 3 mer TAT 22% conserved"	
repeat_region	8371. .8689	
repeat_region	/notes="AluY repeat: matches 1. .307 of consensus"	
repeat_region	8504. .8515	
repeat_region	/notes="12.0 copies 1 mer A 24% conserved"	
repeat_region	8665. .8709	
repeat_region	/note="9.0 copies 5 mer AAAC 44% conserved"	
repeat_region	8666. .8712	
repeat_region	/notes="11.8 copies 4 mer AAC 42% conserved"	
repeat_region	8748. .8763	
repeat_region	/notes="16.0 copies 1 mer T 32% conserved"	
repeat_region	9005. .9122	
repeat_region	/note="MIR repeat: matches 50. .180 of consensus"	
Query Match		
Best Local Similarity 99.5%; Score 2805.8; DB 9; Length 167078;		
Matches 2818; Conservative 0; Mismatches 2; Indels 1; Gaps 1;		
Qy	1	ATCCTAAGCGCACGCTGGGAGCGACACACTGGGGAGTGGTCTGAGGATCTGGGT 60
Db	114296	ATCCTAAGCGCACAGCTGGGAAAGCAGACACTGGGGAAAGTGGTCTGGGGATGTGGGT 114355
Qy	61	CACCTGGGGTCAAGGTGGAGCTTTTCAGGGTCTCCCGTCAATGAGCTGAGTTTCTTTGGC 120
Db	114356	CACCTGGGGTGAAGTGGAGCTTTTCAGGGTCTCCCGTCAATGAGCTGAGTTTCTTTGGC 114415
Qy	121	AGGGAATTTACCAAGCTGAAGAAAGCTGCGCGGAGAGCTCAAACTGAGCAAGGCCAGC 180
Db	114416	AGGGAATTTACCAAGCTGAAGAAAGCTGCGCGGAGAGCTCAAACTGAGCAAGGCCAGC 114475
Qy	181	TGCTCACACCGGAGGAGTGTGGACAGGATCTTCCTCTCGTGGATGAGATGAGATG 240
Db	114476	TGCTCACACCGGAGGAGTGTGGACAGGATCTTCCTCTCGTGGATGAGATGAGATG 114535
Qy	241	GTAAGAGGGCGAGAGATGGGGAGTGTGTCCACTCTGCATCATCGCCACTTTCTGGCC 300
Db	114536	GTAAGAGGGCGAGAGATGGGGAGTGTGTCCACTCTGCATCATCGCCACTTTCTGGCC 114595
Qy	301	GCACGTCTTTGGGCAAGGCCCTCCACCTTCCAAACCTCGGGGTCTCATCTGTGAGAAGC 360
Db	114596	GCACGTCTTTGGGCAAGGCCCTCCACCTTCCAAACCTCGGGGTCTCATCTGTGAGAAGC 114655
Qy	361	TGTGGAGAGATGTCATGAACCTAACAAAGGACTCATGAGCAGTGTGTGTAGGATGAC 420
Db	114656	TGTGGAGAGATGTCATGAACCTAACAAAGGACTCATGAGCAGTGTGTGTAGGATGAC 114715
Qy	421	TAAAGTCTCTACAGGAGTTCTCTATGAGGCCAGGCACGCAGAAATAGAAGAAATAGGAAC 480
Db	114716	TAAAGTCTCTACAGGAGTTCTCTATGAGGCCAGGCACGCAGAAATAGAAGAAATAGGAAC 114775
Qy	481	TTTGGAGTCAGGCAGGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAATTTCTCTATC 540
Db	114776	TTTGGAGTCAGGCAGGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAATTTCTCTATC 114835
Qy	541	TGGAAATGGGATTAATAGTGGTTTGAGAGAAATGAATAGGATTAATGTTTAAAGC 600
Db	114836	TGGAAATGGGATTAATAGTGGTTTGAGAGAAATGAATAGGATTAATGTTTAAAGC 114895

Db 1257 ATCCCAAGACGCACAGCCTGGGAAGCCAGCACTGGGAAGTGTGCTGGGGGATGTGGT 1316
QY 61 CACTGGGTGAAGGTGAGAGCTTTTCAAGGGTCTCCCGTCAATGACGTGAGATTTTCTTTGGC 120
Db 1317 CACTGGGGTGAAGGTGAGAGCTTTTCAAGGGTCTCCCGTCAATGACGTGAGATTTTCTTTTGC 1376
QY 121 AGGGAATTTACCACTGAAGAAAGCCTGCCGCGAGAGCTACAAACTGACGCAAGCCACGC 180
Db 1377 AGGGAATTTACCACTGAAGAAAGCCTGCCGCGAGAGCTACAGACTGAGCAAGACCACGC 1436
QY 181 TGCTCACACCCGAGAGAGTCTGTGACAGAGATCTTCTCTCTGTGTGATGAGAATGAGATG 240
Db 1437 TGCTCACACCCGAGAGAGTCTGTGACAGAGATCTTCTCTCTGTGTGATGAGAATGAGATG 1496
QY 241 GTAAGAGGGGACAGATAGGGGAGAGTGTGTCTCACTCTGCATCATCGCCACTTTCTGGCC 300
Db 1497 GTAAGAGGGGACAGATAGGGGAGAGTGTGTCTCACTCTGCATCATCGCCACTTTCTGGCC 1556
QY 301 GCACGTCTTGGGCAAGCCCTCCACCTTCCAAACCTCGGGTCTCATCTGTGTGAAGGC 360
Db 1557 GCACGTCTTGGGCAAGCCCTCCACCTTCCAAACCTCGGGTCTCATCTGTGTGAAGGC 1616
QY 361 TGTGAGAGAGATGTTCATGAACCTAAACAAAGGGACTCATGAGCACGTGTTTGTAGGAGTGAC 420
Db 1617 TGTGAGAGAGATGTTCATGAACCTAAACAAAGGGACTCATGAGCACGTGTTTGTAGGAGTGAC 1676
QY 421 TAAAGTCTTACAGAGTTGTGATGAGAGCCAGGACGACAGAAATAGAAAGATAGGAAC 480
Db 1677 TAAAGTCTTACAGAGTTGTGATGAGAGCCAGGACGACAGAAATAGAAAGATAGGAAC 1736
QY 481 TTTGAGTTCAGGACGGAGTGATATTAGAGCTTCTCGTCTAGTCTCAATTTCTCTCATC 540
Db 1737 TTTGAGTTCAGGACGGAGTGATATTAGAGCTTCTCGTCTAGTCTCAATTTCTCTCATC 1796
QY 541 TGGAAATGGGGATTAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAGAGC 600
Db 1797 TGGAAATGGGGATTAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAGAGC 1856
QY 601 AGGCATAGGGTATAGCTTCATTTACAGGCTGTCTGGGCTTTTCTCTGTAGGCCAAAGCCC 660
Db 1857 AGGCATAGGGTATAGCTTCATTTACAGGCTGTCTGGGCTTTTCTCTGTAGGCCAAAGCCC 1916
QY 661 AGCCTCAGGCTATCTGGGGAGAGAGCTGGCTTGGGAATACACACTTGAGCCCTCCAGCTC 720
Db 1917 AGCCTCAGGCTATCTGGGGAGAGAGCTGGCTTGGGAATACACACTTGAGCCCTCCAGCTC 1976
QY 721 TCTCAGCTCCACCCAGCAATTTCCGTGGTACCATGCGCAAAAGTAAAACTTCAATTCATCA 780
Db 1977 TCTCAGCTCCACCCAGCAATTTCCGTGGTACCATGCGCAAAAGTAAAACTTCAATTCGTC 2036
QY 781 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTGTGAGATTCAGACACCTGACAGCC 840
Db 2037 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTGTGAGATTCAGACACCTGACAGCC 2096
QY 841 GCAAGCTTGAGCTTGAGACTGACGATGTGGCATAGAGCTGTAGCGCATCTCGG 900
Db 2097 GCAAGCTTGAGCTTGAGACTGACGATGTGGCATAGAGCTGTAGCGCATCTCGG 2156
QY 901 AGCGAGGCTCTCTCTCTCCGCCAGACCCAGGCTCTCCCTTTCTTACATGACACCTC 960
Db 2157 AGCGAGGCTCTCTCTCTCCGCCAGACCCAGGCTCTCCCTTTCTTACATGACACCTC 2216
QY 961 TCTCTCCCTCTCTCAGGCAGCTGTCTGTAACAGATTTGTGAAGGTGCCGCTCGGA 1020
Db 2217 TCTCTCCCTCTCTCAGGCAGCTGTCTGTAACAGATTTGTGAAGGTGCCGCTCGGA 2276
QY 1021 CAAAGTGGGTGAAGATGCTCAGATGACATGAATCCACAGCTGCTCGCTCAGCA 1080
Db 2277 CAAAGTGGGTGAAGATGCTCAGATGACATGAATCCACAGCTGCTCGCTCAGCA 2336
QY 1081 GAGACGGAAAGTGCATGTTCTTGAGGAGTCTGGGGCCCTCTCACGACTCCAGGCTCACC 1140

2337 GAGACGAAAAAGTGCATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCCAGGCTCACC 2396
QY 1141 CAGTTTTCAGGGTAGTAGGAGGGTCCCTGGCTCAGCTGCTCATGCGCACTCTTCCCC 1200
Db 2397 CAGTTTTCAGGGTAGTAGGAGGGT--CCTGGCTCAGCTGCTCATGCGCACTCTTCCCC 2454
QY 1201 TGGTGTTCACATCTCTTGGCACCCCTGTGACGGCTGAGTGGGATGGGGAAGGGCTGCTG 1260
Db 2455 TGGTGTTCACATCTCTTGGCACCCCTGTGACGGCTGAGTGGGATGGGGAAGGGCTGCTG 2514
QY 1261 GTTTGAAGTGCACCAACAGGCAATAGTCCATTTTGGAGAGTCTCTTGGATGGTGAAGG 1320
Db 2515 GTTTGAAGTGCACCAACAGGCAATAGTCCATTTTGGAGGAGTCTCTTGGATGGTGAAGG 2574
QY 1321 AATTCAAGTACTTTTCTTCTTTCAGCCGCTCTTGGAGGACTGTGCTTGGCTGGGTGTT 1380
Db 2575 AATTCAAGTACTTTTCTTCTTTCAGCCGCTCTTGGAGGACTGTGCTTGGCTGGGTGTT 2634
QY 1381 GTGGGGCTCCACAGTCTTCTGGGTGTTCTCAGTTTGGAAAGCAAGCCCAACTGAGGGGTGA 1440
Db 2635 GTGGGGCTCCACAGTCTTCTGGGTGTTCTCAGTTTGGAAAGCAAGCCCAACTGAGGGGTGA 2694
QY 1441 GGGTCCCAACAGCAACCAATCAGAAATAGAAACAAAGACTGGTAGGAGCGAGGGGTGGGA 1500
Db 2695 GGGTCCCAACAGCAACCAATCAGAAATAGAAACAAAGACTGGTAGGAGCGAGGGGTGGGA 2754
QY 1501 GGGTGTTCAGACTGAAGAAAGGACGAGGAGTTGCCGGGACGCTGCTCACGCTCTGAATC 1560
Db 2755 GGGTGTTCAGACTGAAGAAAGGACGAGGAGTTGCCGGGACGCTGCTCATGCTCTGAATC 2814
QY 1561 CCAGCACTTTTGGAGGCGCGGCGGAGATCAAGAGTTCAGGAGATCGAGACCACTCCTG 1620
Db 2815 CCAGCACTTTAGGAGGCGCGGCGGAGATCAAGAGTTCAGGAGATCGAGACCACTCCTG 2874
QY 1621 GCTAACACGGGTGAAACCCCGTCTTCTACTAAAAATACAAAAAATCAGCCGGGTGAGGTG 1680
Db 2875 GCTAACACGGGTGAAAA--CTAATACTAATAAAAAATAAAAAATAAGCCGTGTGTAGGTG 2932
QY 1681 CGGGGCGCTGTAGTCCCGAGCTACTCAGGAGGCTGAGGCAGAGAAATGGCGTGAACCCCA 1740
Db 2933 CGGGGCGCTGTGTATCCAGCTACTCAGGAGGCTGAGGCAGAGATTTAGCGTGAACCCCA 2992
QY 1741 GGGGGC 1746
Db 2993 GGGGCC 2998

RESULT 5
HSM805690
LOCUS HSM805690 2288 bp mRNA linear PRI 16-JUN-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp686E1183 (from clone DKFZp686E1183); complete cds.
ACCESSION BX537393
VERSION BX537393.1 GI:31873287
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2288)
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (16-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMPZ (Biomedical Research Center at the Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686E1183) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further

information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cdnA/>.

```
FEATURES
    source
        1..2288
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /map="6p21.1"
            /clone="DKFZp686E1183"
            /tissue_type="human retina"
            /clone_lib="686 (synonym: h1cc3). Vector pSport1_sfi; host
            DH10B; sites SfiIA + SfiIB"
            /dev_stage="adult"
        1..2288
            /gene="DKFZp686E1183"
            137..739
            /gene="DKFZp686E1183"
            /note="guanylate cyclase-activating protein 2"
            /codon_start=1
            /product="hypothetical protein"
            /protein_id="CAD97635.1"
            /db_xref="GI:31873288"
            /translation="MGQFSEEAAGRIDVAELQEWYKFPVMECPSGTLMFHFPR
            PKVTDDEASQYVSGMRAPDKNGDNTIDFLEYAALNLVLRGLEHKLKRTPLIYD
            KDNGCIDLKLELINVGGIYQKACRRLEQEQQLLTPBEVDRIFLVDENGDDGQ
            LSLNGFVGARDDKVMKMLQMDMNPSSLAQQRKSAMF"
            2261..2266
            /polya_signal
            /gene="DKFZp686E1183"
            2270
            /polya_site
            /gene="DKFZp686E1183"
    ORIGIN
```

```
Query Match      58.0%; Score 1635.2; DB 9; Length 2288;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1659; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY  977  GGCAGCTGCTCTGAAACGAGTTTGTGTAAGTGTCCTCGGACAAAGTGGGTGATGAAG 1036
DB    |||
QY  1037  ATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCAGAGCGGAAAGTGCC 1096
DB    |||
QY  671  ATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCAGAGCGGAAAGTGCC 730
QY  1097  ATGTTCTGAGGAGTCTGGGCGCCCTCCAGCACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1156
DB    |||
QY  731  ATGTTCTGAGGAGTCTGGGCGCCCTCCAGCACTCCAGGCTCACCCAGGTTTCCAGGGTAG 790
QY  1157  TAGGAGGTTCCCTGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1216
DB    |||
QY  791  TAGGAGGTT-CCCTGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 849
QY  1217  GCACCCCTGTGCAGGCTGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 1276
DB    |||
QY  850  GCACCCCTGTGCAGGCTGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 909
QY  1277  CAGGGCATAGTCCATTTTTCAGGAGTCCCTCGGATGGTGAAGGGAATTCAGTTACTTTTC 1336
DB    |||
QY  910  CAGGGCATAGTCCATTTTTCAGGAGTCCCTCGGATGGTGAAGGGAATTCAGTTACTTTTC 969
QY  1337  CTGTTTCAGCCGCTCTCGGGAGGACTGTGCTTTGGCTGGGTGGGTGGGTGGGTGGGTGG 1396
DB    |||
QY  970  CTGTTTCAGCCGCTCTCGGGAGGACTGTGCTTTGGCTGGGTGGGTGGGTGGGTGGGTGG 1029
QY  1397  TTCTGGGTGTTCTCAGTTTGAAGCAAGACCACTGAGGGGTGAGGGTCCCAAGACCA 1456
DB    |||
QY  1030  TTCTGGGTGTTCTCAGTTTGAAGCAAGACCACTGAGGGGTGAGGGTCCCAAGACCA 1089
QY  1457  ATCAGAAATGAGAACACAAAGACTGGTAGGAGCGAGGGGTGGAGGGTCTTCAGACTGAA 1516
DB    |||
QY  1090  ATCAGAAATGAGAACACAAAGACTGGTAGGAGCGAGGGGTGGAGGGTCTTCAGACTGAA 1149
QY  1517  GAAAGGCGAGGAGTTGCGCGGCAACGCTGGCTCAGCGCTGTAATCCAGCACTTTGGGAGG 1576
```

```

DB  1150  GAAAGGCGAGGAGTTGCGGGGCAACGCTGGCTCAGCGCTGTAATCCAGCACTTTGGGAGG 1209
QY  1577  CCGAGCGGGGAGATCACGAGGTGAGGACCATCTCTGGCTAAACGCGGTGAA 1636
DB    |||
DB  1210  CCGAGCGGGGAGATCACGAGGTGAGGACCATCTCTGGCTAAACGCGGTGAA 1269
QY  1637  ACCCGTCTCTACTAAAAATACAAAAATCAGCGGGGTGAGGTGGCGGCGCTGTAGTC 1696
DB    |||
DB  1270  ACCCGTCTCTACTAAAAATACAAAAATCAGCGGGGTGAGGTGGCGGCGCTGTAGTC 1329
QY  1697  CCAGCTACTCAGGAGGCTGAGGCAAGAGNATGGCTGTAACCCAGGGGCGGAGCCTACA 1756
DB    |||
DB  1330  CCAGCTACTCAGGAGGCTGAGGCAAGAGNATGGCTGTAACCCAGGGGCGGAGCCTACA 1389
QY  1757  GTGAGCGGAGATTGCGCCACTGCACTCCAGCCTGACGACAGTGAAGCTCCGCTC-AAA 1815
DB    |||
DB  1390  GTGAGCGGAGATTGCGCCACTGCACTCCAGCCTGACGACAGTGAAGCTCCGCTC-AAA 1449
QY  1816  AAAAAAAGAAAGAAAGAAAGAGAGGAGGAGTTTGGGGGCGAGGGGCGAGCAATAATT 1875
DB    |||
DB  1450  AAAAAAAGAAAGAAAGAAAGAGAGGAGGAGTTTGGGGGCGAGGGGCGAGCAATAATT 1509
QY  1876  CTATAACTTCCGGGATGCTGAGGGGCTTCATGGGGAGGAGACCTGCGCTCTCTCCCTCA 1935
DB    |||
DB  1510  CTATAACTTCCGGGATGCTGAGGGGCTTCATGGGGAGGAGACCTGCGCTCTCTCTCCCTCA 1569
QY  1936  AGGCATCTCACAGTGTGTCAACAGGAAATGCGCAAAATCGCTGCAAGGCTGTGG 1995
DB    |||
DB  1570  AGGCATCTCACAGTGTGTCAACAGGAAATGCGCAAAATGCGCAAAATGCGCTGTGG 1629
QY  1996  TCTTTCTGCTTTGAAAGGCTCAGCTGTACTTTAAAGGAGCTGTTTCAGCTCTGCTCGGT 2055
DB    |||
DB  1630  TCTTTCTGCTTTGAAAGGCTCAGCTGTACTTTAAAGGAGCTGTTTCAGCTCTGCTCGGT 1689
QY  2056  GCTGCTCTGGGACCCCTGCTGCGAACCAACCACTCCCCCAACATCTCTCTTTCATC 2115
DB    |||
DB  1690  GCTGCTCTGGGACCCCTGCTGCGAACCAACCACTCCCCCAACATCTCTCTTTCATC 1749
QY  2116  CATATCCCCAGTATGGACCTTCCCAACTCCAGCATAAGCTGAATGTTTCTCTTTAA 2175
DB    |||
DB  1750  CATATCCCCAGTATGGACCTTCCCAACTCCAGCATAAGCTGAATGTTTCTCTTTAA 1809
QY  2176  AGGATGGAGAAATCTCTGCTGCTGCTGGCAAGAAATGGGGGACTGTTGACTGGGATTG 2235
DB    |||
DB  1810  AGGATGGAGAAATCTCTGCTGCTGCTGGCAAGAAATGGGGGACTGTTGACTGGGATTG 1869
QY  2236  TGGGCTGGGCTTGGCTTCTAACTGCTGTGACCCAGACAGCCACTTCTCTCCCTAAC 2295
DB    |||
DB  1870  TGGGCTGGGCTTGGGCTTCTAACTGCTGTGACCCAGACAGCCACTTCTCTCCCTAAC 1929
QY  2296  CTTGGTTATGCTTTCGCGACAGTGAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2355
DB    |||
DB  1930  CTTGGTTATGCTTTCGCGACAGTGAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 1989
QY  2356  TGTGTTTTAGATGGAAATTAATTTTGTATATAAACTTTATGTGTAAACCCCAATAT 2415
DB    |||
DB  1990  TGTGTTTTAGATGGAAATTAATTTTGTATATAAACTTTATGTGTAAACCCCAATAT 2049
QY  2416  AGAAATAGATTAAGGGAGTCTCTGTTGAAAGGGGAGCTGAGTACCCTCTGGAAAC 2475
DB    |||
DB  2050  AGAAATAGATTAAGGGAGTCTCTGTTGAAAGGGGAGCTGAGTACCCTCTGGAAAC 2109
QY  2476  TGGAGGCACTCTGAAAGAAAGCAAACTGAAACCAAGTGGCTGGGTGACTGTTACTCTTA 2535
DB    |||
DB  2110  TGGAGGCACTCTGAAAGAAAGCAAACTGAAACCAAGTGGCTGGGTGACTGTTACTCTTA 2169
QY  2536  TAAGACAGTTTAAAGTGAAGCTGGAAACCAATTTGCTTTTACCTTTGAATAGATAGT 2595
DB    |||
DB  2170  TAAGACAGTTTAAAGTGAAGCTGGAAACCAATTTGCTTTTACCTTTGAATAGATAGT 2229
QY  2596  TAAGTGGTATATAAGAAATAAAACTAACCTATTAAACCTGAGA 2639
DB    |||
```



```
Db      2230 TATGTTGGTATATAGAAAATAAACTAACTATTATAAAGCTGAAA 2273

RESULT 6
HSM807325/c
LOCUS      HSM807325      2967 bp      mRNA      linear      PRI 30-AUG-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp686A2293 (from clone DKFZp686A2293).
ACCESSION BX647181
VERSION    BX647181.1 GI:34366209
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 2967)
AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
CONSRSTM  The German Human cDNA Consortium
TITLE     Direct Submission
JOURNAL   Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
COMMENT   Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKFZp686A2293) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.

FEATURES             source
    source
        1..2967
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZp686A2293"
            /tissue_type="human retina"
            /clone_lib="686 (synonym: hlec3). Vector pSport1_sfi; host
            DH10B; sites SfiIA + SfiIB"
            /dev_stage="adult"
            polyA_site . 2948
ORIGIN
Query Match      33.9%; Score 956.6; DB 9; Length 2967;
Best Local Similarity 99.3%; Pred. No. 1.9e-258;
Matches 972; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 1532 GCCGGCACGGTGGCTCAGCGCTGTAATCCAGCAGCACTTTGGAGGCCGCGGCGGAGAT 1591
Db 996 GCCGGCGCAGTGGCTCAGCGCTGTCATCCAGCAGCACTTTGGAGGCCGCGGCGGAGAT 937
QY 1592 CACGAGGTCAGGAGATCAGACCATCTGGCTTAACACGGGGTGAACCCCGTCTACTA 1651
Db 936 CACGAGGTCAGGAGATCAGACCATCTGGCTTAACACGGGGTGAACCCCGTCTACTA 877
QY 1652 AAAATACAAAAAATCAGCGGGTGAAGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAG 1711
Db 876 AAAATACAAAAAATCAGCGGGTGAAGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAG 817
QY 1712 GCTGAGGCAAGAGAAATGGCGTGAACCCAGCGGGGCCGAGCCCTACAGTGGCGGAGATTGC 1771
Db 816 GCTGAGGCAAGAGAAATGGCGTGAACCCAGCGGGGCCGAGCCCTACAGTGGCGGAGATTGC 757
QY 1772 GCCACTGCACTCCAGCTGGAGCAGCAGTGAGACTCCGTCTC---AAAAAAGAGAGAA 1828
Db 756 GCCACTGCACTCCAGCTGGAGCAGCAGTGAGACTCCGTCTCAAAAAAAGAGAGAA 697
QY 1829 AGAAAGAGAGAGAGAGAGTTTGGGGGCGAGGGGGCAGCAATAATCTATAACTTCCGG 1888
Db 696 AGAAAGAGAGAGAGAGAGTTTGGGGGCGAGGGGGCAGCAATAATCTATAACTTCCGG 637
QY 1889 GATGCTGAGGGCGGTTTCATGGGAGGAGCCCTGGCGCTCTCTCCCAAGGCAATCCTACC 1948
|||||
```

```
636 GATGCTGAGGGCGGTTTCATGGGAGGAGCCCTGGCGCTCTCTCTCCCAAGGCAATCCTACC 577
1949 AGTGGTGTCAACAGAGAAAATGGCAGCAATAACGCTGCAGGCTGTGGTCTTTCTGCGCTTT 2008
576 AGTGGTGTCAACAGAGAAAATGGCAGCAATAACGCTGCAGGCTGTGGTCTTTCTGCGCTTT 517
2009 GAAAGGGGTACGCTGTACTTTAAAGGGAGCTGTTTTCAGCTCTGCGCTGGGTGCTGCTCTGGGAC 2068
516 GAAAGGGGTACGCTGTACTTTAAAGGGAGCTGTTTTCAGCTCTGCGCTGGGTGCTGCTCTGGGAC 457
2069 CCCCTGTGTCACACCCAGCACTCCCCCAACAATCTCTCTTTCCATTCATATCCCCAGT 2128
456 CCCCTGTGTCACACCCAGCACTCCCCCAACAATCTCTCTTTCCATTCATATCCCCAGT 397
2129 ATGACACCTTCCACAACCTCCAGGCATAGCTGAATGTTTCTCTTTAAAGGATGGAGAAA 2188
396 ATGACACCTTCCACAACCTCCAGGCATAGCTGAATGTTTCTCTTTAAAGGATGGAGAAA 337
2189 CTTCTGTCTGTCTCTGGCAAGAAATGGGGGAGCTGTTGACTGGGATTTGGGCTGGGCTTG 2248
336 CTTCTGTCTGTCTCTGGCAAGAAATGGGGGAGCTGTTGACTGGGATTTGGGCTGGGCTTG 277
2249 GCTTCTAACTGCTGTGTGACCCCAAGACGACCACTTCTCCTCCCTAACTTGGTTATGTCT 2308
276 GCTTCTAACTGCTGTGTGACCCCAAGACGACCACTTCTCCTCCCTAACTTGGTTATGTCT 217
2309 TGGCAGCAGCTGTGAGCAGTGGGAGCTGGGAGCAAGCTGTTGATTTGATTTGTTTAGAT 2368
216 TGGCAGCAGCTGTGAGCAGTGGGAGCTGGGAGCAAGCTGTTGATTTGATTTGTTTAGAT 157
2369 GTGGAAATTTATTTTGTATATAAACTCTTATGTGTAAACCCCAATATAGAACTAGATTA 2428
156 GTGGAAATTTATTTTGTATATAAACTCTTATGTGTAAACCCCAATATAGAACTAGATTA 97
2429 AAAGGGAGTCTCTCTGGTTGAAAGGGAGCTGAGTACCTCTGGAACCTGGAGGCACTCT 2488
96 AAAGGGAGTCTCTCTGGTTGAAAGGGAGCTGAGTACCTCTGGAACCTGGAGGCACTCT 37
2489 GAAAAAGCAAACTGAAA 2507
36 GAAAAAGCAAACTGAAA 18

RESULT 7
G30061/c
LOCUS      human STS SHGC-36091, sequence tagged site.
DEFINITION G30061
ACCESSION  G30061
VERSION     G30061.1 GI:1593612
KEYWORDS   STS; STS sequence; primer; sequence tagged site.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 407)
AUTHORS   Myers,R.M.
JOURNAL   Unpublished (1996)
COMMENT   Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu
            Primer A: CAAGGTAAAGCAAAATGTTTTC
            Primer B: AGGCACCTCTGAAAAAGCA
            STS size: 104
            PCR Profile:
                Initial incubation: 94 degrees C for 90 seconds
                Denaturation:      94 degrees C for 15 seconds
```

Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600
Protocol:
Template: 25 ng
each 1 uM
dNTPs: each 200 uM
Tag Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from H40682
-- Washington University/Merck EST sequence.

FEATURES
source

Location/Qualifiers
1..407
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="6"

STS
primer_bind 38..141
primer_bind 38..60
complement (122..141)

ORIGIN

Query Match 13.3%; Score 374; DB 11; Length 407;
Best Local Similarity 98.0%; Pred. No. 6.3e-94;
Matches 399; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY 2214 GGGGACCTCTGACCTGGGATTTGGGCTTGGCTTCTAAGTCTGTGACCCCAAG 2273
Db 407 GGGGGCTGTGACTGGGATTTGGGCTTGGGCTTCTAAGTCTGTGACCCCAAG 348
QY 2274 ACAGCCACTTCTCCTCCCTAACCTTGGTTATGTCTTGGCAGCACAGTGGAGTGGGAC 2333
Db 347 ACAGCCACTTCTCCTCCCTAACCTTGGTTATGTCTTGGCAGCACAGTGGAGTGGGAC 288
QY 2334 TAGGGACACAGTTTGGATTTATTTGTTTATAGATGGAATTTTGTATATATAA 2393
Db 287 TAGGGACACAGTTTGGATTTATTTGTTTATAGATGGAATTTTGTATATATAA 228
QY 2394 CTCTTATGTATACCCCAATAGAACTAGATTTAAAGGGAGTCTCTCTGTTGAAAGG 2453
Db 227 CTCTTATGTATACCCCAATAGAACTAGATTTAAAGGGAGTCTCTCTGTTGAAAGG 168
QY 2454 GGAG-CTGAGTACCTCTCGAACTGGAGGCACCTCTGAAAAAGCAAACTGAAAAACCACT 2512
Db 167 GGAGCTGAGTACCTCTCGAACTGGAGGCACCTCTGAAAAAGCAAACTGAAAAACCACT 108
QY 2513 GCCCTGGGTCACTGTTACTCTCTATAGA-CAGTTTAAAGTGAGACTTGAAAAACATTGG 2571
Db 107 GCCCTGGGTCACTGTTACTCTCTATAGANCAGTTTAAAGTGAGNCCTGAAAAACATTGG 48
QY 2572 CTTTACCTTGAATAGATAGTGTTTTATGTTGTTATATAGAAATAA 2618
Db 47 CTTTACCTTGAATAGATAGTGTTTTATGTTGTTATATAGCAATAA 1

RESULT 8
CQ491451/c
LOCUS CQ491451 1506 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 23318 from Patent WO0160860.
ACCESSION CQ491451
VERSION CQ491451.1 GI:41457070
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 23318 23-AUG-2001;
FEATURES Millennium Predictive Medicine, Inc. (US)
source Location/Qualifiers
1..1506
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 11.9%; Score 337; DB 6; Length 1506;
Best Local Similarity 99.7%; Pred. No. 2.1e-83;
Matches 348; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 977 GGCCAGCTGTCTCTGAACGAGTTTGTGAAGTGGCCGTCGGGACAAAGTGGGTGATGAAG 1036
Db 1325 GGCCAGCTGTCTCTGAACGAGTTTGTGAAGTGGCCGTCGGGACAAAGTGGGTGATGAAG 1266
QY 1037 ATGCTGCAGATGCATGATCCAGCAGCTGGCTCGCTCAGCAGACGCGAAAAGTGCC 1096
Db 1265 ATGCTGCAGATGCATGATCCAGCAGCTGGCTCGCTCAGCAGACGCGAAAAGTGCC 1206
QY 1097 ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCAGAGGTCCACCCAGGTTTCCAGGTTAG 1156
Db 1205 ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCAGAGGTCCACCCAGGTTTCCAGGTTAG 1146
QY 1157 TAGGAGGTCCCTCGCTCAGCTCTCATGCCCCACTTCTCCCTGGTGTGATTCCTG 1216
Db 1145 TAGGAGGT-CCCTCGCTCAGCTCTCATGCCCCACTTCTCCCTGGTGTGATTCCTG 1087
QY 1217 GCACCCCTCTGCAGGGCTGAGTGGGATGGGGAAGGGCTGCTGGGTTTGAAGTGGCCAA 1276
Db 1086 GCACCCCTCTGCAGGGCTGAGTGGGATGGGGAAGGGCTGCTGGGTTTGAAGTGGCCAA 1027
QY 1277 CAGGCGATAGTCCATTTTGGAGAGTCCCTGGGATGGTGAAGGAATTC 1325
Db 1026 CAGGCGATAGTCCATTTTGGAGAGTCCCTGGGATGGTGAAGGAATTC 978

RESULT 9
CQ497332/c
LOCUS CQ497332 1506 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 29199 from Patent WO0160860.
ACCESSION CQ497332
VERSION CQ497332.1 GI:41462968
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 29199 23-AUG-2001;
FEATURES Millennium Predictive Medicine, Inc. (US)
source Location/Qualifiers
1..1506
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 11.9%; Score 337; DB 6; Length 1506;
Best Local Similarity 99.7%; Pred. No. 2.1e-83;
Matches 348; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 977 GGCCAGCTGTCTCTGAACGAGTTTGTGAAGTGGCCGTCGGGACAAAGTGGGTGATGAAG 1036
Db 1325 GGCCAGCTGTCTCTGAACGAGTTTGTGAAGTGGCCGTCGGGACAAAGTGGGTGATGAAG 1266


```

misc_feature 66101..69100
              /note="assembly_fragment:06949"
misc_feature 69201..72200
              /note="assembly_fragment:06950"
misc_feature 72301..75300
              /note="assembly_fragment:06951"
misc_feature 75401..78400
              /note="assembly_fragment:06952"
misc_feature 78501..81500
              /note="assembly_fragment:06953"
misc_feature 81601..84600
              /note="assembly_fragment:06954"
misc_feature 84701..87700
              /note="assembly_fragment:06955"
misc_feature 87801..90800
              /note="assembly_fragment:06956"
misc_feature 90901..93900
              /note="assembly_fragment:06957"
misc_feature 94001..97000
              /note="assembly_fragment:06958"
misc_feature 97101..105500
              /note="assembly_fragment:06961"
misc_feature 105601..108600
              /note="assembly_fragment:06962"
misc_feature 108701..1140438
              /note="assembly_fragment:06983"
misc_feature 140539..333201
              /note="assembly_fragment:06982
              clone_end:T7
              vector_side:right"

ORIGIN
Query Match      9.9%; Score 278.6; DB 2; Length 333201;
Best Local Similarity 83.5%; Pred. No. 1.1e-66;
Matches 340; Conservative 0; Mismatches 64; Indels 3; Gaps 2;

QY 1515 AAAAAAGCAGGAGTTCGGGGGCGCGGTGAGTTCAGCCCTGTAATCCAGCAGCTTTGGGA 1574
Db 99432 AAAATTACAAAAAATTGGCGGGCGCGGTGAGTTCAGCCCTGTAATCCAGCAGCTTTGGGA 99491
QY 1575 GCGCCAGCGGGCAGATCAGCAGGTCAGGAGATCAGACCATCTGGCTTAACACGGGGTG 1634
Db 99492 GCGCCAGCGGGCGGATCAGCAGGTCAGGAGATCAGACCATCTGGCTTAACAC--GGTG 99549
QY 1635 AAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGGTGAGTTCGGGGCGCGCTGTAG 1694
Db 99550 AAACCCGTCTCTACTAAAAATACAAAAATTAGCGGGGAGGTGGCGGGCGCTGTAG 99609
QY 1695 TCCAGCTACTCAGAGGCTGAGGCAAGAGATGCGTGAACCCAGGGGGCGGAGCCTA 1754
Db 99610 TCCAGCTACTCCGAGGCTGAGGCGAGGAGATGCGTGAACCCAGGGGGTGGAGCCTG 99669
QY 1755 CAGTACGCGAGATTGCCCACTGCACTCCAGCTGAGGAGCAGTGAAGTCCGCTCA 1814
Db 99670 CAGTACGCGAGATTGCCCACTGCACTCCAGCTGCGGCGACAGCAGACTCCGCTCA 99729
QY 1815 AAAAAAAGCAAGAAAAAGAGGAGGAGTTCGGGGCGAGGGGCGAGCAATAT 1874
Db 99730 AAAAAAAGCAAGAAAAAATACAAAAATTAAGCCGGGATGTCGGGGCGCTGTAGT 99789
QY 1875 TCTATACTTCGGGATGTCGAGGGCGGTTCATGGGAGGACCTGG 1921
Db 99790 CCAGCTACT-CCGAGGCTGAGGCGAGGAGATGCGGTGAACCCAGG 99835

RESULT 11
AC023158/c 146530 bp DNA linear PRI 01-APR-2003
LOCUS Homo sapiens 12 BAC RP11-438D14 (Roswell Park Cancer Institute
DEFINITION Human BAC Library) complete sequence.
ACCESSION AC023158
VERSION AC023158.28 GI:11496286
KEYWORDS HTG.

```

SOURCE ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 146530)

Muzny,D.M., Adams,C.,

Albrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbaria,J.,

Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,

Bowle,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,

Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,

Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,

Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,

Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,

Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,

Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,

Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,

Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,

Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,

Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,

Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,

Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O.,

Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Honsi,F.,

Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,

Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

Karissom,B., Kelly,S., Khan,U., King,L., Korvan,J., Kovar,C.,

Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,

Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,

Loulseghe,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,

Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,

Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P.,

Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,

Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,

Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,

Nguyen,S., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,

Ogih,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,

Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,

Quiles,M., Ren,Y., Rives,M., Rojbas,A., Rojubokan,I., Rolfe,M.,

Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shm,C.,

Shooshkar,N., Sisson,I., Sodergren,E., Sonaie,T., Sparks,A.,

Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,

Tanerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,

Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,

Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,

Washington,C., Watlington,S., Williams,G., Williamson,A.,

Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Zhou,J.,

Zorrilla,S., Kucherlapati,R. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 146530)

Worley,K.C.

Direct Submission

Submitted (09-FEB-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 146530)

Worley,K.C.

Direct Submission

Submitted (01-DEC-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 146530)

Worley,K.C.

Direct Submission

Submitted (08-FEB-2001) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 146530)

Worley,K.C.

Direct Submission

Submitted (01-APR-2003) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Dec 1, 2000 this sequence version replaced gi:11192041.

COMMENT

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

```
----- Summary Statistics -----
Contig length: 191218
Phrap values in estimate: 190513
Average error rate (BCW-Phrap estimate): 3.84429e-05
Fraction of Phrap values less than 40 : 0.0103405
Number of consensus changing edits: 15
Number of N's in consensus : 0
```

```
----- Consensus changing edits -----
Position Original+Context Edited+Context
5392 ttcgactaa(n)attagttaga ttcgactaa(a)attagttaga
72712 actgcaagct(n)tgctctccgg actgcaagct(c)tgctctccgg
72764 agtagctggg(g)ctacagccac agtagctggg(a)ctacagccac
72780 ggcaccacc(c)ccccaggcc ggcaccacc(a)ccccaggcc
72783 accacncc(c)ccaggctgg accacncc(a)ccaggctgg
72794 ccaggctgg(n)taaatctttt ccaggctgg(c)taaatctttt
72799 cctggncaaa(a)ttttttttt cctggncaaa(t)ttttttttt
72938 gccaccgtgc(n)cagcaaaaag gccaccgtgc(c)cagcaaaaag
108090 ttttttga(n)ggagctctac ttttttga(t)ggagctctac
108125 tggagtgcag(g)ggcacaatct tggagtgcag(t)ggcacaatct
108136 ggcacaatct(n)gntcactgc ggcacaatct(c)gntcactgc
108139 acaatcngg(n)tcactgcaac acaatcngg(c)tcactgcaac
159377 ccttttaac(n)tttggctgtt ccttttaac(c)tttggctgtt
181705 ttttgggtgg(n)aaaaaaaaa ttttgggtgg(a)aaaaaaaaa
181755 gctcttttta(n)cccttacatg gctcttttta(a)cccttacatg
```

----- Distribution of Quality < 40 Bases -----

```
1000 |
900 |
800 |
```

```
# bases
700 |
600 |
500 |
400 |
300 |
200 |
100 |
0 |
```

Phrap Value Range

FEATURES

```
Version: 1.01 qxfo
Location/Qualifiers
1. .146530
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clones="RP11-438D14"
misc_feature
1. .6343
/note="overlaps bases 80628..86970 of clone AC016956"
repeat_region
694..795
/rpt_family="L2"
repeat_region
complement(796..1097)
/rpt_family="AluSp"
repeat_region
1098..1124
/rpt_family="L2"
STS
1417..1566
/standard_name="G07737"
/db_xref="dbSTS:13583"
repeat_region
complement(1615..1709)
/rpt_family="MIR"
repeat_region
complement(3095..3160)
/rpt_family="L2"
repeat_region
3317..3419
/rpt_family="MER104"
repeat_region
4052..4088
/rpt_family="AT_rich"
repeat_region
4220..4695
/rpt_family="MLT1J"
repeat_region
4922..4950
/rpt_family="AT_rich"
repeat_region
5402..5439
/rpt_family="(CAGA)n"
```

```
Query Match 9.9%; Score 278; DB 9; Length 146530;
Best Local Similarity 87.1%; Pred. No. 1.6e-66;
Matches 317; Conservative 0; Mismatches 45; Indels 2; Gaps 1;
QY 1488 GCGAGGGTGGAGGGTGTTCAGACTGAAGAAAGCGAGGAGTTGCCGGGACGGTGGCT 1547
DB 89495 GCGAGGTGCGAACAAGGGTTATGTCAAGGTAAATTTATCTAGGCGGGCGGGTGGCT 89436
QY 1548 CACGCTGTATCCAGCACCTTTGGGAGCGCGGCGGAGATCACCAGGTCAGGAGAT 1607
DB 89435 CACGCTGTATCCAGCACCTTTGGGAGCGCGGCGGAGATCACCAGGTCAGGAGAT 89376
QY 1608 CGAGACCATCTGGGTAAACAGGGGTGAACCCCGTCTCTACTAAAAATACAAAAATCA 1667
DB 89375 CGAGACCATCTGGGTAAACAGGGGTGAACCCCGTCTCTACTAAAAATACAAAAATCA 89318
QY 1668 GCCGGGTGAGTGGCGGGCGCTGTAGTCCAGCTACTCAGAGGCTCAGGCAAGAGAT 1727
DB 89317 GCCGGGTGAGTGGCGGGCGCTGTAGTCCAGCTACTCAGAGGCTCAGGCAAGAGAT 89258
QY 1728 GCGGTGAACCCAGGGGCGCGGCTACAGTGAGCGCGGAGATTGCGCCACTGCACCTCCAGC 1787
DB 89257 GCGGTGAACCCAGGGGCGCGGCTACAGTGAGCGCGGAGATTGCGCCACTGCACCTCCAGC 89198
```



```

/Chromosome="6"
/clone="DAAP-200B17"
/clone_lib="APD"

ORIGIN
Query Match          9.8%; Score 277.8; DB 9; Length 78827;
Best Local Similarity 88.9%; Pred. No. 1.7e-66;
Matches 312; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 1515 AAGAAAAGCAGGAGTTGCCGGGACAGGTGGCTACGGCTGTAAATCCAGCAGCTTTGGGA 1574
Db 51750 AAAAAAAAAAAAAAAAAAGCGCGCGGTGGCTACGGCTGTAAATCCAGCAGCTTTGGGA 51809

QY 1575 GCGCAGCGCGGAGTCACAGAGTCAGAGATCGAGACCATCTGGCTTAACACAGCGGTG 1634
Db 51810 GCGCAGCGCGGTGGATCAGAGTCAGAGATCGAGACCATCTGGCTTAACAC--GGTG 51867

QY 1635 AAACCCCGTCTCTACTAAAAATACAAAAATACAGCCGGGTGAGGTGGCGGCGCTGTAG 1694
Db 51868 AAACCCCGTCTCTACTAAAAATACAAAAATTTAGCGGGCGAGGTGGCGGCGCTGTAG 51927

QY 1695 TCCAGCTACTCAGAGGCTGAGGCAAGAGATGGCGTGAACCCAGCGGGCGAGCCTA 1754
Db 51928 TCCTAGCTACTCGGAGGCTGAGGAGGAGATGGCGTGAACCCAGGAGCGGAGCCTG 51987

QY 1755 CAGTGAGCGGAGATGGCCACTGCACTCCAGCTTGAAGAGTGAAGTCCGCTCTCAA 1814
Db 51988 CAGTGAGCGGAGATGGCCACTGCACTCCAGCTTGAAGAGTGAAGTCCGCTCTCAA 52047

QY 1815 AAAAAAAAAAGAAAGAAAGAAAGGAGGAGTGTGGGGGCGAGGGGCG 1865
Db 52048 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATTTGGGACAGATGTGGC 52098

RESULT 14
CR388382/c
LOCUS Human DNA sequence from clone DADB-23P23 on chromosome 6, complete
DEFINITION sequence.
ACCESSION CR388382
VERSION CR388382.4 GI:47678787
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Brown,J.
Direct Submission
Submitted (24-JUL-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 25, 2004 this sequence version replaced gi:47604289.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
DADB-23P23 is from a DNA-arts DBB human bac library VECTOR:
pBelosAC11.

FEATURES
Source
Location/Qualifiers
1..109067
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DADB-23P23"
/clone_lib="DNA-arts-BAC.1-DBB.1"

ORIGIN
Query Match          9.8%; Score 277.8; DB 9; Length 109067;
Best Local Similarity 88.9%; Pred. No. 1.7e-66;
Matches 312; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 1515 AAGAAAAGCAGGAGTTGCCGGGACAGGTGGCTACGGCTGTAAATCCAGCAGCTTTGGGA 1574
Db 82323 AAAAAAAAAAAAAAAAAAGCGCGCGGTGGCTACGGCTGTAAATCCAGCAGCTTTGGGA 82264

QY 1575 GCGCAGCGCGGAGATCAGAGGTCAGAGATCGAGACCATCTGGCTTAACACAGCGGTG 1634
Db 82263 GCGCAGCGCGGTGGATCAGAGGTCAGAGATCGAGACCATCTGGCTTAACAC--GGTG 82206

QY 1635 AAACCCCGTCTCTACTAAAAATACAAAAATTTAGCGGGCGAGGTGGCGGCGCTGTAG 1694
Db 82205 AAACCCCGTCTCTACTAAAAATACAAAAATTTAGCGGGCGAGGTGGCGGCGCTGTAG 82146

QY 1695 TCCAGCTACTCAGAGGCTGAGGCAAGAGATGGCGTGAACCCAGCGGGCGAGCCTA 1754
Db 82145 TCCTAGCTACTCGGAGGCTGAGGAGGAGATGGCGTGAACCCAGGAGCGGAGCCTG 82086

QY 1755 CAGTGAGCGGAGATGGCGGCTGCACTCCAGCTTGAAGAGTGAAGTCCGCTCTCAA 1814
Db 82085 CAGTGAGCGGAGATGGCGGCTGCACTCCAGCTTGAAGAGTGAAGTCCGCTCTCAA 82026

QY 1815 AAAAAAAAAAGAAAGAAAGAAAGGAGGAGTGTGGGGGCGAGGGGCG 1865
Db 82025 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATTTGGGACAGATGTGGC 81975

RESULT 15
CR3759926/c
LOCUS Homo sapiens chromosome 6 clone DAAP-269H15, *** SEQUENCING IN
DEFINITION PROGRESS ***, 13 unordered pieces.
ACCESSION CR3759926
VERSION CR3759926.1 GI:51965193
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Sims,S.
Direct Submission
Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

```


----- Project Information
Center project name: BAP269H15
----- Summary Statistics
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 143896 bases at least Q40
Consensus quality: 144892 bases at least Q30
Consensus quality: 145693 bases at least Q20
Insert size: 147081; sum-of-contigs
Quality coverage: 4.63x in Q20 bases; sum-of-contigs Quality
coverage: 8.94x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2610: contig of 2610 bp in length
2611 2710: gap of 100 bp
2711 10514: contig of 7804 bp in length
10515 10814: gap of 100 bp
10815 13221: contig of 2607 bp in length
13222 13321: gap of 100 bp
13322 16471: contig of 3150 bp in length
16472 16571: gap of 100 bp
16572 24423: contig of 7852 bp in length
24424 24523: gap of 100 bp
24524 26634: contig of 2111 bp in length
26635 26734: gap of 100 bp
26735 31590: contig of 4856 bp in length
31591 31690: gap of 100 bp
31691 42950: contig of 11260 bp in length
42951 43050: gap of 100 bp
43051 66753: contig of 23703 bp in length
66754 66853: gap of 100 bp
66854 81357: contig of 14504 bp in length
81358 81457: gap of 100 bp
81458 90446: contig of 8989 bp in length
90447 90546: gap of 100 bp
90547 110259: contig of 19713 bp in length
110260 110359: gap of 100 bp
110360 148281: contig of 37922 bp in length.

FEATURES

Location/Qualifiers
1..148281
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DAP-269H15"
/clone_lib="APD"
1..2610
/note="assembly fragment:00039
fragment chain:1"
2711..10514
/note="assembly fragment:00186
fragment chain:1"
10615..13221
/note="assembly fragment:00063
fragment chain:1"
13322..16471
/note="assembly fragment:00079
fragment chain:2"
16572..24423
/note="assembly fragment:00129
fragment chain:2"
24524..26634
/note="assembly fragment:00029
fragment chain:2"
26735..31590

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature
/note="assembly fragment:00102
fragment chain:3"
31691..42950
/note="assembly fragment:00244
fragment chain:3"
43051..66753
/note="assembly fragment:00685
fragment chain:4"
66854..81357
/note="assembly fragment:00417
fragment chain:4"
81458..90446
/note="assembly fragment:00325.0"
90547..110259
/note="assembly fragment:00519.0"
110360..148281
/note="assembly fragment:00854.0"

ORIGIN

Query Match 9.8%; Score 277.8; DB 2; Length 148281;
Best Local Similarity 88.9%; Pred. No. 1.8e-66;
Matches 312; Conservative 0; Mismatches 37; Indels 2; Gaps 1;
QY 1515 AAGRAAGCGGAGCTTGGCGGCACGGTGGCTCAGCGCTGAATCCAGCAGCTTTGGGA 1574
DB 129689 AAAAAAAAAAAAAAGGCGCGGCTGCGCTGTAATCCAGCAGCTTTGGGA 129630
QY 1575 GGCCGAGCGGGCAGATCACGAGGTCAGGAGATCGAGACCATCTGGCTTAACACCGGGGTG 1634
DB 129629 GGCCGAGCGGGTGGATCACGAGGTCAGGAGATCGAGACCATCTGGCTTAACAC 129572
QY 1635 AAACCCCGTCTCTACTAAAAATACAAAAATACAGCCGGGTGAGGTGGCGGGCGCTGTAG 1694
DB 129571 AAACCCCGTCTCTACTAAAAATACAAAAATATAGCGGCGAGGTGGCGGGCGCTGTAG 129512
QY 1695 TCCAGCTACTCAGGAGGCTGAGGCAAGAGATGCGTGAACCCAGGGGGCGGAGCCTTA 1754
DB 129511 TCCTAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCAGGAGGGGAGCCTG 129452
QY 1755 CAGTGAGCGGAGATTGCGCCACTGCACCTCCAGCCTGGAGCAGACAGTGAGACTCCGTCTCAA 1814
DB 129451 CAGTGAGCGGAGATTGCGCCACTGCACCTCCAGCCTGGGCGACAGCGAGACTCCGTCTCAA 129392
QY 1815 AAAAAAAAAAGAAAGAAAGAAAGGAGGAGTTTGGGGGGCGAGGGGGC 1865
DB 129391 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTGGGACAGATGTGGC 129341

Search completed: June 29, 2005, 01:06:02
Job time : 8111.93 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:28:52 ; Search time 987.126 Seconds
(without alignments)
16917.381 Million cell updates/sec

Title: US-08-731-499-6
Perfect score: 2821
Sequence: 1 ATCTTAAGACGACACGCTG.....AGGGTTCAAGACGACGCTG 2821

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2808.6	99.6	2820	AAV09020	Homo sapi
2	337	11.9	1506	ABV29181	Human pro
3	337	11.9	1506	ABV23329	Human pro
4	325.8	11.5	425	AAI83630	Human pol
5	275	9.7	32152	AAK89020	Human dig
6	275	9.7	32152	AAK91534	Human dig
7	275	9.7	32152	AAI57791	Human col
8	275	9.7	32152	AAI57791	Human col
9	275	9.7	32152	AAI57791	Human col
10	275	9.7	32152	ABV23329	Human pro
11	275	9.7	32152	ABV23329	Human pro
12	275	9.7	32152	ABV23329	Human pro
13	275	9.7	32152	ABV23329	Human pro
14	274.6	9.7	335	AAI99076	Human exc
15	274.6	9.7	335	AAI99076	Human exc
16	273.4	9.7	238484	AAI63426	Human kid
17	272.6	9.7	7995	AAI63426	Human kid
18	272.6	9.7	54378	AAI63426	Human kid
19	271.8	9.6	45000	AAI63426	Human kid
20	271.8	9.6	122186	AAI63426	Human kid

C 21	271	9.6	3540	4	AAK72282	Human imm
C 22	270	9.6	810	4	AAK72282	Human imm
C 23	269.2	9.5	70469	11	AAK72282	Human imm
C 24	268.6	9.5	600	4	AAK72282	Human imm
C 25	266.8	9.5	15270	4	AAK72282	Human imm
C 26	266.8	9.5	15270	4	AAK72282	Human imm
C 27	266.4	9.4	51469	4	AAK72282	Human imm
C 28	266.4	9.4	51469	4	AAK72282	Human imm
C 29	266.4	9.4	51469	4	AAK72282	Human imm
C 30	266.4	9.4	51469	4	AAK72282	Human imm
C 31	263.6	9.3	816	4	AAK72282	Human imm
C 32	263.6	9.3	816	4	AAK72282	Human imm
C 33	263.6	9.3	816	4	AAK72282	Human imm
C 34	259.8	9.2	68495	13	AAK72282	Human imm
C 35	259	9.2	110000	9	AAK72282	Human imm
C 36	259	9.2	110000	12	AAK72282	Human imm
C 37	258.6	9.2	31304	4	AAK72282	Human imm
C 38	258.6	9.2	31304	4	AAK72282	Human imm
C 39	258.6	9.2	31304	4	AAK72282	Human imm
C 40	258.6	9.2	31304	4	AAK72282	Human imm
C 41	258.6	9.2	31304	4	AAK72282	Human imm
C 42	258.6	9.2	31304	4	AAK72282	Human imm
C 43	258.6	9.2	31304	4	AAK72282	Human imm
C 44	257.4	9.1	15306	4	AAK72282	Human imm
C 45	257.2	9.1	57248	6	AAK72282	Human imm

ALIGNMENTS

RESULT 1
AAV09020
ID AAV09020 standard; cDNA; 2820 BP.
XX
AC AAV09020;
XX
DT 21-JUL-1998 (first entry)
XX
DE Homo sapiens 20q13 amplicon GCAP transcript.
XX
KW 20q13 amplicon; chromosome 20; tumour; detection; cyclic AMP; cAMP;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentosa;
KW Leber's congenital amaurosis; biosynthesis;
KW guanine cyclase activating protein; ds.
OS Homo sapiens.
XX
PN WO9802539-A1.
XX
PD 22-JAN-1998.
XX
PF 15-JUL-1997; 97WO-US012343.
XX
PR 15-JUL-1996; 96US-00680395.
PR 16-OCT-1996; 96US-00731499.
PR 17-JAN-1997; 97US-00785532.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX WPI; 1998-110587/10.
XX
PT New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
PT treating diseases.
XX
PS Claim 1; Page 63; 91pp; English.
XX
CC The sequence is that of a cDNA sequence GCAP, which was isolated from the
CC 20q13 amplicon. It encodes a guanine cyclase activating protein which is
CC involved in the biosynthesis of cyclic AMP. Sequences from this gene can

be used in the treatment of retinal degeneration. It can be used as a probe for the detection of chromosomal abnormalities at 20q13. It and other sequences isolated from the 20q13 amplicon are consistently amplified in primary tumours. These sequences are useful as probes or as probe targets for monitoring the relative copy number of corresponding sequences from a biological sample such as tumour cells. The sequences can also be used in therapeutic applications for modulating the expression of the endogenous gene or the activity of the gene product. Examples of therapeutic approaches include antisense inhibition of gene expression, gene therapy, and monoclonal antibodies that specifically bind the gene products. The products can also be used in the treatment of other diseases, e.g. age-related macular degeneration, Leber's congenital amaurosis and retinitis pigmentata

```
Query Match          99.6%; Score 2808.6; DB %; Length 2820;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2819; Conservative 1; Mismatches 0; Indels 1; Gaps 1
```

Qy	1	ATCTTAAGACGCACAGCCTCTGGGAAGCCAGCACTCTGGGAAGTGGTGTGATGAGGATGTGGGT	60
Db	1	ATCTTAAGACGCACAGCCTCTGGGAAGCCAGCACTGGGAAGTGGTGTGATGAGGATGTGGGT	60
Qy	61	CACCTGGGGTGAAGTGGAGCTTTCAAGGTCTCCCGTCAATGCAGCTGAGTTTCTTTTCGC	120
Db	61	CACCTGGGGTGAAGTGGAGCTTTCAAGGTCTCCCGTCAATGCAGCTGAGTTTCTTTTCGC	120
Qy	121	AGGGAATTTACCGCTGAAGAAAGCCTCGCGCGAGAGCTACAAACTGAGCAAGGCCAGC	180
Db	121	AGGGAATTTACCGCTGAAGAAAGCCTCGCGCGAGAGCTACAAACTGAGCAAGGCCAGC	180
Qy	181	TGCTCAACCCGAGAGTCTGTGGACAGGATCTTCTCTCTGTGTGATGAGAAATGAGATG	240
Db	181	TGCTCAACCCGAGAGTCTGTGGACAGGATCTTCTCTCTGTGTGATGAGAAATGAGATG	240
Qy	241	GTAAGAGGGCAGAGATGGGAGAGTGTCTCCACTCTGCATCATCGCCACCTTTCTGGCC	300
Db	241	GTAAGAGGGCAGAGATGGGAGAGTGTCTCCACTCTGCATCATCGCCACCTTTCTGGCC	300
Qy	301	GCACGTCTCTGGGCAAGGCCCTCCACCTTCCAAACCTCGGGTCTCATCTGTGAAAGGC	360
Db	301	GCACGTCTCTGGGCAAGGCCCTCCACCTTCCAAACCTCGGGTCTCATCTGTGAAAGGC	360
Qy	361	TGTGGAGAGATGTCAATGAATTAACAAAGGAGCTCATGAGCAGCTGTTGTATGAGATGAC	420
Db	361	TGTGGAGAGATGTCAATGAATTAACAAAGGAGCTCATGAGCAGCTGTTGTATGAGATGAC	420
Qy	421	TAAAGTCTCAGGAGTTGCTGATGGAGGCCAGGCAACGAGAAATAGGAATAGGAAC	480
Db	421	TAAAGTCTCAGGAGTTGCTGATGGAGGCCAGGCAACGAGAAATAGGAATAGGAAC	480
Qy	481	TTTGGAGTCAGGCAGGGAGTGATATATTAGGCTTCTCGTCTTAGTCTCAATTTCTCTATC	540
Db	481	TTTGGAGTCAGGCAGGGAGTGATATATTAGGCTTCTCGTCTTAGTCTCAATTTCTCTATC	540
Qy	541	TGGAAATGGGGATTAATATAGTGTGAGAGGAATGAATAGGATTAATGTGTTTAAGAGC	600
Db	541	TGGAAATGGGGATTAATATAGTGTGAGAGGAATGAATAGGATTAATGTGTTTAAGAGC	600
Qy	601	AGGCATAGGGTAGACCTCCATTTCAGGCTGCTTGGGCTTTTCCTCCCTGTAGCCCAAGGCC	660
Db	601	AGGCATAGGGTAGACCTCCATTTCAGGCTGCTTGGGCTTTTCCTCCCTGTAGCCCAAGGCC	660
Qy	661	AGCCTCAGGGCTATGTGGGGAGAGCTGGCTTGGAAATACACATTTGAGCCCTCAGGCTC	720
Db	661	AGCCTCAGGGCTATGTGGGGAGAGCTGGCTTGGAAATACACATTTGAGCCCTCAGGCTC	720
Qy	721	TCTCAGCTCCACCCAGCATTTCCGTGGTAGCATCGGCAAAAGTAACCTTCAATTCATCA	780
Db	721	TCTCAGCTCCACCCAGCATTTCCGTGGTAGCATCGGCAAAAGTAACCTTCAATTCATCA	780
Qy	781	GCAAGAAAGCCCTTTAAAGGTGGCAGGAGACTCTCTGGAGATTTCAGACACCTTCAGAGCC	840

PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX Disclosure; SEQ ID NO 2596; 986pp; English.
XX The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention
XX
SQ Sequence 32152 BP; 9229 A; 7256 C; 7335 G; 8332 T; 0 U; 0 Other;
Query Match 9.7%; Score 275; DB 4; Length 32152;
Best Local Similarity 88.1%; Pred. No. 7, 5e-60;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;
QY 1486 GAGGCGGGTGGAGGCTTTGAGACTGAGAAAGCAGGAGTTGCCGGGCACGGTGG 1545
DB 13869 GAGGCGGGTGGATCATGAAGTGAAGAGACCAATCCTGGTTGGCGCGCGGTGG 13928

QY 1546 CTCACGCTGTAATCCACGACACTTTGGGAGGCCGAGGCGGCGAGATCACGAGTTCAGGAG 1605
DB 13929 CTCACGCTGTAATCCACGACACTTTGGGAGGCCGAGGCGGCGAGATCACGAGTTCAGGAG 13988
QY 1606 ATCGAGACCATCTCGGTGCTAAACACGGGGTGAAACCCCGTCTTCTACTTAAATAACAAAAAT 1665
DB 13989 ATCGAGACCATCTCGGTGCTAACAC--GGTGAACCCCGTCTTCTACTTAAATAACAAAAAT 14046
QY 1666 CAGCGGGTGAGGTGGCGGGCGCCTGTAGTCCAGTACTCAGGAGGCTGAGGCAAGAGA 1725
DB 14047 TAGCCGGCGAGGTGGCGGGCGCCTGTAGTCCAGTACTCAGGAGGCTGAGGCAAGAGA 14106
QY 1726 ATGGCGTCAACCCCGAGGGCGGAGCCTACAGTGAAGCCGAGATTGCGGCACCTGCCTCCA 1785
DB 14107 ATGGCGTCAACCCCGAGGGCGGAGCCTGAGTGAAGCCGAGATTGCGGCACCTGCCTCCA 14166
QY 1786 GCCTGGACGACAGTGAGACTCCGCTCTCAAAAAAAGAAAAAGAAAAAGAAA 1838
DB 14167 GCCTGGCGCACAGCGAGACTCCGCTCTCAAAAAAAGAAAAAAGAAAAAGAAA 14219
RESULT 6
AAK91534/c
ID AAK91534 standard; DNA; 32152 BP.
XX
AC AAK91534;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen genomic sequence SEQ ID NO: 5110.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001324.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246603P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-502630/55.
XX
XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
PS Disclosure; SEQ ID NO 5110; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention
XX
SQ Sequence 32152 BP; 8332 A; 7335 C; 7256 G; 9229 T; 0 U; 0 Other;
Query Match 9.7%; Score 275; DB 4; Length 32152;
Best Local Similarity 88.1%; Pred. No. 7, 5e-60;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;
QY 1486 GAGGCGGGTGGGAGGGTGTGAGACTGAGAGAAAGGAGGAGGTGCGCGGCACGGTGG 1545
Db 18284 GAGGCGGGTGGATCATGAAGTGAAGATCAAGACCATCTGTGTCGGCGCGGGTGG 18225
QY 1546 CTCAGCCTGTAAATCCAGCAGCTTTGGGAGCGCGGAGGCGGAGATCAGAGGTGAGGAG 1605
Db 18224 CTCAGCCTGTAAATCCAGCAGCTTTGGGAGCGCGGAGGCGGAGTACAGGTGAGGAG 18165
QY 1606 ATCGAGACCATCTCGCTAAACACCGGGTGAACACCCCGCTCTTACTTAAATAACAAAAAT 1665
Db 18164 ATCGAGACCATCTCGCTTAAACAC--GGTGAACCCCGCTCTTACTTAAATAACAAAAAT 18107
QY 1666 CAGCCGGGTGAGGTGGCGGGCGCCTGTAGTCCAGCTACTCAGGAGGTGAGGCAAGAGA 1725

```
Db 18106 TAGCCGGCGAGGTGGCGGCGCTGTAGTCCACCTACTCGGAGGCTGAGGCAGGAGA 18047
Qy 1726 ATGGCGTGAACCCAGCGGGCGAGCCCTACAGTGAGCGAGATTGCGGCACCTGCCTCCA 1785
Db 18046 ATGGCGTGAACCCAGCGGGCGAGCCCTGCAGTGAGCGAGATTGCGGCACCTGCCTCCA 17987
Qy 1786 GCCTGGACGACGAGTCCGCTCTCAAAAAAAAAAAGAAAGAAA 1838
Db 17986 GCCTGGCGGACGAGACTCCGCTCTCAAAAAAAAAAAGAAAAGAA 17934

RESULT 7
AAI57791/C
ID AAI57791 standard; DNA; 32152 BP.
XX
AC AAI57791;
XX
DT 19-OCT-2001 (first entry)
XX
DE Human colorectal cancer antigen coding sequence SEQ ID NO: 328.
XX
KW Human; colorectal cancer; colorectal cancer antigen; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155350-A1.
XX
PD
XX
PF 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001350.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0218290P.
PR 14-AUG-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0223287P.
PR 01-SEP-2000; 2000US-0223343P.
PR 01-SEP-2000; 2000US-0223344P.
PR 01-SEP-2000; 2000US-0223345P.
PR 05-SEP-2000; 2000US-0223509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
```


PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 13-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239353P.
PR 13-OCT-2000; 2000US-0239373P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-02411785P.
PR 20-OCT-2000; 2000US-02411786P.
PR 20-OCT-2000; 2000US-02411787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246533P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.

```
Db      14167 GCCTGGGACAGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAGA 14219
RESULT 9
ID      ABS99968/c
XX      ABS99968 standard; DNA; 32152 BP.
AC      ABS99968;
XX      18-DEC-2002 (first entry)
DT      Genomic DNA #172 encoding human colorectal cancer related protein.
XX      Human; colorectal cancer related protein; colon; rectum;
XX      colorectal cancer metastasis; gastrointestinal disorder; cytostatic;
XX      gene; ds.
OS      Homo sapiens.
XX      US2002119919-A1.
PN      29-AUG-2002.
XX      17-JAN-2001; 2001US-00764855.
PF      31-JAN-2000; 2000US-0179065P.
XX      (ROSE/) ROSEN C A.
PA      (RUBE/) RUBEN S M.
PA      (BARA/) BARASH S C.
XX      Rosen CA, Ruben SM, Barash SC;
PI      WPI; 2002-731367/79.
XX      New colorectal cancer polypeptide for diagnosing, prognosing, preventing,
PT      and treating immune, hyperproliferative, liver, kidney, reproductive
PT      disorders and for identifying modulators of therapeutic use.
XX      Disclosure; SEQ ID NO 328; 183pp; English.
XX      The present invention relates to the isolation of novel human colorectal
CC      cancer related proteins, and polynucleotide sequences encoding them. The
CC      sequences of the invention are useful in the diagnosis, treatment,
CC      prevention and/or prognosis of the colon and/or rectum, including
CC      colorectal cancer, colorectal cancer metastases, and gastrointestinal
CC      disorders such as dysphagia, peptic esophagitis, gastric reflux,
CC      irritable bowel syndrome, and peritoneal diseases. The invention also
CC      describes antibodies that bind colorectal cancer related proteins,
CC      vectors, host cells, and recombinant and synthetic methods for producing
CC      human colorectal cancer related polynucleotides, polypeptides, and/or
CC      antibodies. ABS99797-ABS99974 represent genomic sequences encoding human
CC      colorectal cancer related proteins. Note: The sequence data for this
CC      patent did not form part of the printed specification, but was obtained
CC      in electronic format directly from the USPTO web site at
CC      seqdata.uspto.gov/psipsidentry.html
XX      Sequence 32152 BP; 8332 A; 7335 C; 7256 G; 9229 T; 0 U; 0 Other;
SQ      Query Match      9.7%; Score 275; DB 6; Length 32152;
      Best Local Similarity 88.1%; Pred. No. 7.5e-60;
      Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;
      1486 GAGGCGGGGTGGAGGGGTGTGAGACTGAACAAAGGCGGAGTTGCCGGGCACGGTGG 1545
      18284 GAGGCGGGGTGGATCATGATGAGAGATCAAGACCATCTCTGGTGGCGGGCGGGTGG 18225
      1546 CTCACGCTGTATCCACGACTTTTGGGAGGCGGCGGCGGAGATCAGAGTTCAGGAG 1605
      18224 CTCACGCTGTATCCACGACTTTTGGGAGGCGGCGGCGGAGTTCAGAGTTCAGGAG 18165
      1606 ATCGAGACCATCTCGGTAAACCGGGGTGAACCCCGTCTCTACTAAAAATACAAAAAT 1665
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      18164 ATCGAGACCATCTCGGTAAACAC--GGTGAAACCCCGTCTCTACTAAAAATACAAAAAT 18107
QY      1666 CAGCCGGGTGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGA 1725
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18106 TAGCCGGGCGAGGTGGCGGGCGCTGTAGTCCAGCTACTCGGAGGCTGAGGCAAGAGA 18047
QY      1726 ATGGCGTGAAACCCAGGGGGCGGAGCCTACAGTGAGCGGAGATTGGCCCACTGCCTCCA 1785
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18046 ATGGCGTGAAACCCAGGGGGCGGAGCCTGCAGTGAGCGGAGATTGGCCCACTGCCTCCA 17987
QY      1786 GCCTGAGCAGACGTGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAGAAAAA 1838
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17986 GCCTGGGACAGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAGAAAAA 17934
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
RESULT 10
ADB32581
ID      ADB32581 standard; DNA; 32152 BP.
XX      ADB32581;
AC      ADB32581;
XX      04-DEC-2003 (first entry)
DT      Human novel colon related polypeptide DNA SEQ ID NO 518.
XX      gene therapy; cancer; liver disorder; hepatitis; neural disorder;
XX      Alzheimer's disease; human; colon; ds.
OS      Homo sapiens.
XX      US2003050231-A1.
PN      13-MAR-2003.
XX      17-JAN-2001; 2001US-00764872.
XX      31-JAN-2000; 2000US-0179065P.
PR      04-FEB-2000; 2000US-0180628P.
PR      24-FEB-2000; 2000US-0184664P.
PR      02-MAR-2000; 2000US-0186350P.
PR      16-MAR-2000; 2000US-0189874P.
PR      17-MAR-2000; 2000US-0190076P.
PR      18-APR-2000; 2000US-0198123P.
PR      19-MAY-2000; 2000US-0205515P.
PR      27-JUN-2000; 2000US-0209467P.
PR      28-JUN-2000; 2000US-0214886P.
PR      30-JUN-2000; 2000US-0215135P.
PR      07-JUL-2000; 2000US-0216647P.
PR      07-JUL-2000; 2000US-0216880P.
PR      11-JUL-2000; 2000US-0217487P.
PR      14-JUL-2000; 2000US-0218290P.
PR      26-JUL-2000; 2000US-0220963P.
PR      26-JUL-2000; 2000US-0220964P.
PR      14-AUG-2000; 2000US-0224518P.
PR      14-AUG-2000; 2000US-0224519P.
PR      14-AUG-2000; 2000US-0225213P.
PR      14-AUG-2000; 2000US-0225214P.
PR      14-AUG-2000; 2000US-0225266P.
PR      14-AUG-2000; 2000US-0225267P.
PR      14-AUG-2000; 2000US-0225268P.
PR      14-AUG-2000; 2000US-0225270P.
PR      14-AUG-2000; 2000US-0225447P.
PR      14-AUG-2000; 2000US-0225757P.
PR      14-AUG-2000; 2000US-0225758P.
PR      14-AUG-2000; 2000US-0225759P.
PR      18-AUG-2000; 2000US-0226279P.
PR      22-AUG-2000; 2000US-0226681P.
PR      22-AUG-2000; 2000US-0226689P.
PR      22-AUG-2000; 2000US-0227182P.
PR      23-AUG-2000; 2000US-0227009P.
PR      30-AUG-2000; 2000US-0228924P.
PR      01-SEP-2000; 2000US-0229287P.
```


Db 14107 ATGGCGTGAACCCAGGGGCGGAGCCTGTGACGTGAGCCGAGATTGGCCACTGCACTCCA 14166
QV 1786 GCGTCGACGACAGTGCAGACTCCGCTCTCAAAAAAAGAAAAAGAAAAAGAAA 1838
Db 14167 GCGTCGCGACAGCGAGACTCCGCTCTCAAAAAAAGAAAAAAGAAAAAAGA 14219

RESULT 11
ADB93121/c
ID ADB93121 standard; DNA; 32152 BP.
XX
AC ADB93121;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human colorectal cancer related polypeptide DNA #172.
XX
KW ds; gene; human; colorectal cancer; antigen; gene therapy;
KW gastrointestinal disorder; inflammatory disease; infection; cancer;
KW intestinal neoplasm; small intestine carcinoma tumour;
KW small intestine non-Hodgkin's lymphoma; small bowel lymphoma; ulcer;
KW peptic ulcer; Bruton's disease; X linked infantile agammaglobulinaemia;
KW severe combined immunodeficiency; DiGeorge anomaly;
KW hyperproliferative disorder; acute lymphoblastic leukaemia;
KW acute lymphocytic leukaemia; urinary system disorder; cortical necrosis;
KW kidney infection; cardiovascular disorder; carcinoma heart disease;
KW arrhythmia; respiratory disorder; non-allergic rhinitis; sinusitis;
KW musculoskeletal system disorder; Albers-Schonberg disease;
KW Marfan's syndrome; neurological disease; phenylketonuria;
KW Wernicke's encephalopathy; Alzheimer's disease; endocrine disorder;
KW Grave's disease; Cushing's syndrome; reproductive system disorder;
KW prostatitis; benign prostatic hypertrophy; benign prostatic hyperplasia;
KW thrombosis; atherosclerosis; myocardial infarction; ischaemic attack.
OS Homo sapiens.
XX
XX US2003054420-A1.
XX
XX 20-MAR-2003.
XX
XX 11-FEB-2002; 2002US-00072349.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 07-JUN-2000; 2000US-0205515P.
XX 28-JUN-2000; 2000US-0209467P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.

18-AUG-2000; 2000US-0226279P.
22-AUG-2000; 2000US-0226681P.
22-AUG-2000; 2000US-0226868P.
23-AUG-2000; 2000US-0227182P.
23-AUG-2000; 2000US-0227009P.
30-AUG-2000; 2000US-022824P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
05-SEP-2000; 2000US-0229513P.
06-SEP-2000; 2000US-0230437P.
06-SEP-2000; 2000US-0230438P.
08-SEP-2000; 2000US-0231242P.
08-SEP-2000; 2000US-0231243P.
08-SEP-2000; 2000US-0231244P.
08-SEP-2000; 2000US-0231413P.
08-SEP-2000; 2000US-0231414P.
08-SEP-2000; 2000US-0232080P.
08-SEP-2000; 2000US-0232081P.
12-SEP-2000; 2000US-0231968P.
14-SEP-2000; 2000US-0232397P.
14-SEP-2000; 2000US-0232398P.
14-SEP-2000; 2000US-0232399P.
14-SEP-2000; 2000US-0232400P.
14-SEP-2000; 2000US-0232401P.
14-SEP-2000; 2000US-0233063P.
14-SEP-2000; 2000US-0233064P.
14-SEP-2000; 2000US-0233065P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234997P.
25-SEP-2000; 2000US-0234998P.
26-SEP-2000; 2000US-0235484P.
27-SEP-2000; 2000US-0235834P.
27-SEP-2000; 2000US-0235836P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239935P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241221P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241786P.
20-OCT-2000; 2000US-0241787P.
20-OCT-2000; 2000US-0241808P.
20-OCT-2000; 2000US-0241809P.
01-NOV-2000; 2000US-0244617P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764855.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI, 2003-708345/67.
XX
XX Novel colorectal cancer antigen useful for treating, preventing,
PT diagnosing and/or prognosing gastrointestinal disorders, infections,
PT cancers such as intestinal neoplasms, ulcers.
XX
XX Disclosure; SEQ ID NO 328; 179pp; English.
XX
XX The invention relates to a colorectal cancer antigen. The antigen is
CC useful for chromosome identification, chromosome mapping, radiation
CC hybrid mapping or gene therapy, or as hybridisation probes for
CC differential identification of the tissues or cell types present in a
CC biological sample. The antigen is useful for treating, preventing,
CC diagnosing and/or prognosing gastrointestinal disorders, including
CC inflammatory diseases and/or conditions, infections, cancers (e.g.
CC intestinal neoplasms (carcinoid tumour of the small intestine, non-
CC Hodgkin's lymphoma of the small intestine, small bowel lymphoma) and
CC ulcers (e.g. peptic ulcers). The antigen and its nucleic acid are useful
CC to provide immunological probes for differential identification of the
CC tissue. The antigen and its nucleic acid are useful for treating,
CC preventing, diagnosing and/or prognosing diseases, disorders and/or
CC conditions of the immune system e.g. Bruton's disease, X linked infantile
CC agammaglobulinaemia, severe combined immunodeficiencies, DiGeorge
CC anomaly, etc. The antigen and its nucleic acid is useful for treating,
CC preventing and/or diagnosing hyperproliferative disorders (e.g. acute
CC lymphoblastic leukaemia, acute lymphocytic leukaemia, etc), urinary

Query Match 9.7%; Score 275; DB 10; Length 32152;
Best Local Similarity 88.1%; Pred. No. 7.5e-60;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;
Qy 1486 GAGGCGGGTGGAGGCTTGAAGTGAAGAACAGGAGGATGCGGGCAGCGTGG 1545
Db 18284 GAGGCGGGTGGATCATGAAGTGAAGAGATCAAGACCATCCTGGTTGGCGGGCGGTGG 18225

Qy 1546 CTCACGCTGTAATCCAGCACATTTGGAGGCCGAGGGCGGAGATCACGAGGTGAGGAG 1605
Db 18224 CTCACGCTGTAATCCAGCACATTTGGAGGCCGAGGGCGGCGGATCACGAGGTGAGGAG 18165
Qy 1606 ATCGAGACCATCTCTGGCTTAACACGGGGTGAACCCCGTCTCTACTTAAATAATACAAAAAT 1665
Db 18164 ATCGAGACCATCTCTGGCTTAACAC--GGTGAACCCCGTCTCTACTTAAATAATACAAAAAT 18107
Qy 1666 CAGCCGGGTGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGA 1725
Db 18106 TAGCCGGCGGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGA 18047
Qy 1726 ATGGCGTGAACCCAGAGGGGCGGAGCCCTACAGTGAGCGGAGATTGCGCCACTGCACTCCA 1785
Db 18046 ATGGCGTGAACCCAGAGGGGCGGAGCCCTGCAGTGAGCGGAGATTGCGCCACTGCACTCCA 17987
Qy 1786 GCCTGGACGACGAGTGAAGTCTCCGTCTCAAAAAAAGAAAAAGAAAAAGAA 1838
Db 17986 GCCTGGCGCACAGCGAGACTCCGTCTCAAAAAAAGAAAAAAGAAAAAGAA 17934
RESULT 12
AAK66362
ID AAK66362 standard; DNA; 36933 BP.
XX
XX AAK66362;
XX
XX 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21174.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225575P.
XX 14-AUG-2000; 2000US-0225756P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0233080P.
 PR 08-SEP-2000; 2000US-0233081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234977P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.
 DR
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure; SEQ ID NO 21174; 3071pp + Sequence Listing; English.
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients' own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 36933 BP; 10500 A; 8526 C; 8341 G; 9566 T; 0 U; 0 Other;
 Query Match 9.7%; Score 275; DB 4; Length 36933;
 Best Local Similarity 88.1%; Pred. No. 8e-60;
 Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;
 QY 1486 GAGCGGGTGGGAGGGTGTGACACTGAAGAAAGGAGGAGTTCGCGGCACGGTGG 1545
 Db 13869 GAGCGGGTGGGATCATGATGAAGTGAAGATCAAGACCATCTCTGTTGGCGGCGGGTGG 13928
 QY 1546 CTCACGCTGTAAATCCAGCACTTTGGGAGGCCGAGGCGGCAGATCAGAGGTGAGGAG 1605

Db 13929 CTCAGCCCTGTAATCCAGCACTTTGGAGGCGCGAGGCTCAGAGGTGAGGAG 13988
Qy 1606 ATCGAGACATCTGGCTAAACACGGGGTGAACCCCGCTCTACTATAAATAACAAAAAT 1665
Db 13989 ATCGAGACATCTGGCTAAACAC--GGTGAACCCCGCTCTACTATAAATAACAAAAAT 14046
Qy 1666 CAGCCGGGTGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAGGAGA 1725
Db 14047 TAGCCGGCGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAGGAGA 14106
Qy 1726 ATGGCGTGAACCCACAGGGGGCGGAGCCTACAGTGAGCCGAGATTGGCGCACTGCCA 1785
Db 14107 ATGGCGTGAACCCACAGGGGGCGGAGCCTGCACTGAGCGAGATTGGCGCACTGCCA 14166
Qy 1786 GCCTGGACAGATGAGTACGCTCTCAAAAAAAGAAAAAGAAAAAGAAA 1838
Db 14167 GCCTGGCGACAGCGAGACTCGCTCTCAAAAAAAGAAAAAAGAAAAAGAAA 14219

RESULT 13
ACN44674
ID ACN44674 standard; DNA; 129042 BP.
XX AC ACN44674;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG27276.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
FN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1240; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 129042 BP; 35555 A; 29165 C; 28632 G; 35690 T; 0 U; 0 Other;
Query Match 9.7%; Score 275; DB 11; Length 129042;
Best Local Similarity 88.1%; Pred. No. 1.5e-59;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

Qy 1486 GAGCAGGGGTGGAGGGTGTGAGACTGAAGAAAAAGCAGAGTTGCCGCGCACCGGTGG 1545
Db 123352 GAGCGGGTGGATCATGAGTGAAGAGATCAAGACCATCTCTGGTTGGCGGGCGGTGG 123411
Qy 1546 CTCAGCCCTGTAATCCAGCACTTTGGAGGCCGAGGGCGGAGATCAAGAGGTGAGGAG 1605
Db 123412 CTCAGCCCTGTAATCCAGCACTTTGGAGGCCGAGGGCGGAGATCAAGAGGTGAGGAG 123471
Qy 1606 ATCGAGACCATCTGGCTTAACACGGGGTGAACCCCGCTCTACTATAAATAACAAAAAT 1665
Db 123472 ATCGAGACCATCTGGCTTAACAC--GGTGAACCCCGCTCTACTATAAATAACAAAAAT 123529
Qy 1666 CAGCCGGGTGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAGGAGA 1725
Db 123530 TAGCCGGCGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAGGAGA 123589
Qy 1726 ATGGCGTGAACCCACAGGGGGCGGAGCCTACAGTGAGCCGAGATTGGCGCACTGCCA 1785
Db 123590 ATGGCGTGAACCCACAGGGGGCGGAGCCTGCACTGAGCGAGATTGGCGCACTGCCA 123649
Qy 1786 GCCTGGACAGATGAGTACGCTCTCAAAAAAAGAAAAAGAAAAAGAAA 1838
Db 123650 GCCTGGCGACAGCGAGACTCGCTCTCAAAAAAAGAAAAAAGAAAAAGAAA 123702
RESULT 14
AAI99076/c
ID AAI99076 standard; DNA; 335 BP.
XX AC AAI99076;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human excretory related polynucleotide SEQ ID NO 840.
XX
KW Human; nootropic; neuroprotective; cyostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antiskilling; antianaemic; antithratic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW excretory system; ds.
XX
OS Homo sapiens.
XX
PN WO200155313-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001323.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.

PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226686P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231124P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231124P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231124P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231141P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231141P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	06-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-02559678P.
PR	25-SEP-2000;	2000US-0234998P.	XX		
PR	26-SEP-2000;	2000US-0235484P.	XX		
PR	27-SEP-2000;	2000US-0235834P.	XX		
PR	27-SEP-2000;	2000US-0235836P.	XX		
PR	29-SEP-2000;	2000US-0236327P.	XX		
PR	29-SEP-2000;	2000US-0236367P.	XX		
PR	29-SEP-2000;	2000US-0236368P.	XX		
PR	29-SEP-2000;	2000US-0236369P.	XX		
PR	29-SEP-2000;	20			

```

CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 335 BP; 48 A; 107 C; 82 G; 98 T; 0 U; 0 Other;

Query Match          9.7%; Score 274.6; DB 4; Length 335;
Best Local Similarity 92.0%; Pred. No. 1.1e-60;
Matches 301; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

QY 1513 TGAAGAAAAGCAGAGATTGGCGGCGCAGCGTGGCTCAGCCCTGTATATCCAGCACTTTGG 1572
Db 334 TTATAAAATGCGAGGAAAGCGCGCGGTGGCTCAGCCCTGTATATCCAGCACTTTGG 275

QY 1573 GAGGCCGAGCGGCGCAGATCAGAGGTCAGAGATCGAGACCATCTGGCTAAACACGCGG 1632
Db 274 GAGGCCGAGCGGCGGTCGATCAGAGGTCAGAGATCGAGACCATCTGGCTTAACAC--GG 217

QY 1633 TGAACCCCGTCTCTACTATAAAATCAAAAAATCAGCGGGTGAGGTGGCGGCGCTTGT 1692
Db 216 TGAACCCCGTCTCTACTATAAAATCAAAAAATTAAGCGGCGAGGTGGCGGCGCTTGT 157

QY 1693 AGTCCAGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGGGCGGAGCC 1752
Db 156 AGTCCAGCTACTCGGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGGGCGGAGCC 97

QY 1753 TACAGTGAGCGGAGATTGCGGCACCTGCACCTCGAGCTGGACACAGTGCAGCTCCGCTCTC 1812
Db 96 TGCAGTGAGCGGAGATTGCGGCACCTGCACCTCGAGCTGGGCGACAGCGAGACTCGCTCTC 37

QY 1813 AAAAAAAGAAAGAAAGAAAGAA 1839
Db 36 AAAAAAAGAAAGAAAGAAAGAAAGAA 10

RESULT 15
AAI63426/c
ID AAI63426 standard; DNA; 335 BP.
XX
AC AAI63426;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human kidney related polynucleotide SEQ ID NO 741.
XX
KW Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neotropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antitumor; vulnery; anticonvulsant; antiparasitic;
KW gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; ds.
XX
OS Homo sapiens.
XX
FN WO200155323-A2.
XX
PD 02-AUG-2001.
XX
FF 17-JAN-2001; 2001WO-US001343.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0226279P.
PR 14-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.

```

01-NOV-2000; 2000US-0244617P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249264P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249298P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254077P.
05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-488784/53.
New isolated nucleic acids and polypeptides, useful for diagnosing,
treating and/or preventing human diseases and disorders.
Disclosure; SEQ ID NO 741; 564pp + Sequence Listing; English.
The invention relates to novel kidney related polynucleotides (AAI62971-
AAI63793) and the encoded polypeptides (AAM42417-AM42691) collectively
known as kidney antigens and the use of such kidney antigens for
detecting disorders of the kidney, especially kidney cancer and kidney
cancer metastases. The polynucleotides and proteins are also useful for
preventing, treating or ameliorating medical conditions e.g. by protein
or gene therapy. The genes are isolated from a range of human tissues
disclosed in the specification. The nucleic acids, proteins, antibodies
and (ant)agonists are useful in the diagnosis, treatment and prevention
of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the
adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
lung, or urogenital; (b) immune disorders e.g. Addison's disease,

CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIFO at
XX ftp.wifo.int/pub/published_pct_sequences
SQ Sequence 335 BP; 48 A; 107 C; 82 G; 98 T; 0 U; 0 Other;
Query Match 9.7%; Score 274.6; DB 5; Length 335;
Best Local Similarity 92.0%; Pred. No. 1.1e-60;
Matches 301; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
QY 1513 TGAAGAAAAGCAGGAGTTGCGGGCACGGTGGCTCAGCCCTGTAATCCACACCTTTGG 1572
DB 334 TTATAAATGCGAGGAAGCGCGCGGTGGCTCAGCCCTGTAATCCACACCTTTGG 275
QY 1573 GAGCCGAGCGGGCAGATCAGAGGTGAGGATCGAGACCATCTGCTTAACACGGGG 1632
DB 274 GAGCCGAGCGGGTGGATCAGAGGTGAGGATCGAGACCATCTGCTTAACAC--GG 217
QY 1633 TGAACACCCGCTCTTACTATAAAATACAAAATACAGCGGTGAGGTGGCGGGCGCTGT 1692
DB 216 TGAACACCCGCTCTTACTATAAAATACAAAATATAGCGGGCGAGGTGGCGGGCGCTGT 157
QY 1693 AGTCCAGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGGGGCGAGCC 1752
DB 156 AGTCCAGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGGGGCGAGCC 97
QY 1753 TACAGTGAGCGAGATTGCGCCACTGCACTCCAGCTCGACGACAGTGAAGACTCGGTCTC 1812
DB 96 TCGAGTGAGCGAGATTGCGCCACTGCACTCCAGCTCGGCGACAGAGACTCCGTCTC 37
QY 1813 AAAAAAAGAAAGAAAGAAAGAAAA 1839
DB 36 AAAAAAAGAAAGAAAGAAAGAAAA 10
Search completed: June 28, 2005, 19:33:53
Job time : 994.126 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 18:22:52 ; Search time 320.226 Seconds
(without alignments)
14414.621 Million cell updates/sec

Title: US-08-731-499-6

Perfect score: 2821

Sequence: 1 ATCTAAGACGACAGCCTG.....AGGGGTTCAAGACGACGCTG 2821

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 81813859 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2821	100.0	2821	2	US-08-680-395-6
2	2821	100.0	2821	4	US-08-892-693-6
3	270.8	9.6	601	4	US-09-949-016-17473
4	270.6	9.6	255679	4	US-09-949-016-17189
5	269.6	9.6	51508	4	US-09-949-016-15681
6	269.4	9.5	131332	4	US-09-949-016-15535
7	267.4	9.5	258775	4	US-09-949-016-15435
8	267.2	9.5	24166	4	US-09-949-016-13755
9	267.2	9.5	91665	4	US-09-949-016-12234
10	266.6	9.5	44166	4	US-09-949-016-15829
11	266.6	9.5	54180	4	US-09-949-016-14894
12	266.4	9.4	67479	4	US-09-949-016-11804
13	266.4	9.4	71119	4	US-09-949-016-15358
14	263.6	9.3	86273	4	US-09-949-016-15273
15	263	9.3	360470	4	US-09-949-016-13173
16	262.8	9.3	18048	4	US-09-949-016-12363
17	262.8	9.3	18049	4	US-09-949-016-15235
18	262.8	9.3	136917	4	US-09-949-016-15369
19	262.4	9.3	601	4	US-09-949-016-29990
20	262.4	9.3	601	4	US-09-949-016-160457
21	259.8	9.2	52496	4	US-09-949-016-16118
22	259.8	9.2	52496	4	US-09-949-016-16119
23	259.6	9.2	462589	4	US-09-949-016-12900
24	259.6	9.2	476044	4	US-09-949-016-12412
25	259.2	9.2	601	4	US-09-949-016-40745
26	257.4	9.1	601	4	US-09-949-016-91528
27	256.2	9.1	15725	4	US-09-949-016-17607

c 28	256.2	9.1	106199	4	US-09-949-016-12393	Sequence 12393, A
c 29	256.2	9.1	107980	4	US-09-949-016-14370	Sequence 14370, A
c 30	256	9.1	77100	4	US-09-949-016-16418	Sequence 16418, A
c 31	256	9.1	87863	4	US-09-949-016-14402	Sequence 14402, A
c 32	255.8	9.1	601	4	US-09-949-016-91529	Sequence 91529, A
c 33	255	9.0	84558	4	US-09-949-016-15752	Sequence 15752, A
c 34	254.6	9.0	19091	4	US-09-949-016-15805	Sequence 15805, A
c 35	254.2	9.0	601	4	US-09-949-016-143877	Sequence 143877, A
36	253.6	9.0	78846	4	US-09-949-016-12396	Sequence 12396, A
37	253.6	9.0	78846	4	US-09-949-016-12791	Sequence 12791, A
38	253.6	9.0	78846	4	US-09-949-016-12792	Sequence 12792, A
39	253.6	9.0	78846	4	US-09-949-016-12793	Sequence 12793, A
40	253.6	9.0	78850	4	US-09-949-016-16013	Sequence 16013, A
41	253.6	9.0	78850	4	US-09-949-016-16014	Sequence 16014, A
42	253.6	9.0	78850	4	US-09-949-016-16015	Sequence 16015, A
43	253.6	9.0	78850	4	US-09-949-016-16016	Sequence 16016, A
44	253.6	9.0	78850	4	US-09-949-016-16201	Sequence 16201, A
45	253.6	9.0	78850	4	US-09-949-016-16202	Sequence 16202, A

ALIGNMENTS

RESULT 1
US-08-680-395-6
; Sequence 6, Application US/08680395
; Patent No. 5892010
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W.
; APPLICANT: Collins, Colin
; APPLICANT: Hwang, Soo-in
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowbel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: Genes from the 20q13 Amplicon and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,395
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-06890005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2821 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..2821
; OTHER INFORMATION: /note= "cdna clone GCAP encodes a
; OTHER INFORMATION: guano cyase activating protein
; OTHER INFORMATION:

US-08-680-395-6									
Query Match 100.0%; Score 2821; DB 2; Length 2821; Best Local Similarity 100.0%; Pred. No. 0; Matches 2821; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ATCCTTAAGACGCA	CAGCCTGGGAAGCCAGAC	CTGGGGAAGTGGTCTC	TGAGGATGGGGT	60			
Db	1	ATCCTTAAGACGCA	CAGCCTGGGAAGCCAGAC	CTGGGGAAGTGGTCTC	TGAGGATGGGGT	60			
Qy	61	CACCTGGGGTGAAG	TGGAGCTTTTCAAGGTCT	CCCGTCAATG	CAGCTGATTTCTTTGGC	120			
Db	61	CACCTGGGGTGAAG	TGGAGCTTTTCAAGGTCT	CCCGTCAATG	CAGCTGATTTCTTTGGC	120			
Qy	121	AGGGATTTTACAG	CTGAAGAAAGCCTGCGG	CGAGAGCTACAACT	GAGCAAGGCCAGC	180			
Db	121	AGGGATTTTACAG	CTGAAGAAAGCCTGCGG	CGAGAGCTACAACT	GAGCAAGGCCAGC	180			
Qy	181	TGCTCACACCCG	AGAGGTCTGTGACAGAT	CTTCTCCTGCTG	GATGAGAAATGGAGATG	240			
Db	181	TGCTCACACCCG	AGAGGTCTGTGACAGAT	CTTCTCCTGCTG	GATGAGAAATGGAGATG	240			
Qy	241	GTAAGAGGGG	CAGAGATGGGAGAGT	GTGTCCACTCTG	CAATCGCACTTTCTGGCC	300			
Db	241	GTAAGAGGGG	CAGAGATGGGAGAGT	GTGTCCACTCTG	CAATCGCACTTTCTGGCC	300			
Qy	301	GCAGTCTCTGGG	CAAGGCCCTCCACCTT	CCAACTGGGGTCCCT	CACTCTGTGAAGGC	360			
Db	301	GCAGTCTCTGGG	CAAGGCCCTCCACCTT	CCAACTGGGGTCCCT	CACTCTGTGAAGGC	360			
Qy	361	TGTGAGAAATGT	CATGAACTTAA	CAAAAGGACTCAT	GAGCACGTGTTGTAGGAGTGAC	420			
Db	361	TGTGAGAAATGT	CATGAACTTAA	CAAAAGGACTCAT	GAGCACGTGTTGTAGGAGTGAC	420			
Qy	421	TAAAGTCCTA	CAGAGTTGCTGAT	TGGAGGCGAGGCA	CGCAGAAATAGAAAGATAGGAAAC	480			
Db	421	TAAAGTCCTA	CAGAGTTGCTGAT	TGGAGGCGAGGCA	CGCAGAAATAGAAAGATAGGAAAC	480			
Qy	481	TTTGGAGTCAGG	AGGAGTGATATAT	TGAGCTTCTCGT	CTAGTCTCAATTTCTCATC	540			
Db	481	TTTGGAGTCAGG	AGGAGTGATATAT	TGAGCTTCTCGT	CTAGTCTCAATTTCTCATC	540			
Qy	541	TGGAATAATGGG	GAATAATAGTGT	TGAGAGGAATGAAT	AGGATAATGTGTTTAAAGGC	600			
Db	541	TGGAATAATGGG	GAATAATAGTGT	TGAGAGGAATGAAT	AGGATAATGTGTTTAAAGGC	600			
Qy	601	AGGCATAGGGT	AGACCTCAATTCAGG	CTGTGTGGCTTTCT	CCCTGTAGCCCAAGCCC	660			
Db	601	AGGCATAGGGT	AGACCTCAATTCAGG	CTGTGTGGCTTTCT	CCCTGTAGCCCAAGCCC	660			
Qy	661	AGCCTCAGGGCT	ATGTGGGAGAGAGCT	GGCTTGGGAATACAC	ATTGAGCCCTCCAGCTC	720			
Db	661	AGCCTCAGGGCT	ATGTGGGAGAGAGCT	GGCTTGGGAATACAC	ATTGAGCCCTCCAGCTC	720			
Qy	721	TCTCAGCTCCAC	CCAGCATTTCCGTGGT	ACCATGGCAAAAGTAA	AACTTCAATTCATCA	780			
Db	721	TCTCAGCTCCAC	CCAGCATTTCCGTGGT	ACCATGGCAAAAGTAA	AACTTCAATTCATCA	780			
Qy	781	GCAAGAAAGCCC	TTAAAGGTGGCAGGAG	ACTCTCTGAGATTC	AGACCTGACAAGCC	840			
Db	781	GCAAGAAAGCCC	TTAAAGGTGGCAGGAG	ACTCTCTGAGATTC	AGACCTGACAAGCC	840			
Qy	841	GCAAGCTTGAGG	CTGAGCTGACGAT	TGTGGCATAGACGT	GATAGCGCATCTCTGGG	900			
Db	841	GCAAGCTTGAGG	CTGAGCTGACGAT	TGTGGCATAGACGT	GATAGCGCATCTCTGGG	900			
Qy	901	AGCAGAGTCTCT	CTCTCTGCCCCCAGAC	CCAGGTCTCTTCT	TACATGACCACTC	960			
Db	901	AGCAGAGTCTCT	CTCTCTGCCCCCAGAC	CCAGGTCTCTTCT	TACATGACCACTC	960			
Qy	961	TCCTCCCCCTTG	CTCAGGCGACGTGT	CTGTAA	CGAGTTTGTGAAGTGC	1020			
Db	961	TCCTCCCCCTTG	CTCAGGCGACGTGT	CTGTAA	CGAGTTTGTGAAGTGC	1020			

Qy	2101	TCCTCTCTTCCATCCATATCCCCAGTAGTGGACCTTCCACAACCTCCAGGCCATAAGCTG	2160
Db	2101	TCCTCTCTTCCATCCATATCCCCAGTAGTGGACCTTCCACAACCTCCAGGCCATAAGCTG	2160
Qy	2161	AATGTTTCTCTTTAAAGGATGGAGAAACTTCTCTGTCTCTGGCAAGAAATTTGGGGGAC	2220
Db	2161	AAATGTTTCTCTTTAAAGGATGGAGAAACTTCTCTGTCTCTGGCAAGAAATTTGGGGGAC	2220
Qy	2221	TGTTGACTGGGAATGTGGGCTGGGCTTGGGCTTAACTGCTGTGTGTGACCCAAAGCAGCCCA	2280
Db	2221	TGTTGACTGGGAATGTGGGCTGGGCTTGGGCTTAACTGCTGTGTGTGACCCAAAGCAGCCCA	2280
Qy	2281	CTTCTCTCCCTAAACCTTGGTTATGTCTTGGCAGCAGTGGAGCAGTCCGACTAGGCGGA	2340
Db	2281	CTTCTCTCCCTAAACCTTGGTTATGTCTTGGCAGCAGTGGAGCAGTCCGACTAGGCGGA	2340
Qy	2341	ACAGTTTGGGAATTAATGTGTTTTAGATGTGGAATATTTTTTGTGTATATAAACTCTTAT	2400
Db	2341	ACAGTTTGGGAATTAATGTGTTTTAGATGTGGAATATTTTTTGTGTATATAAACTCTTAT	2400
Qy	2401	GTGTAAACCCCAATATAGAAACTAGATTAAAGGGAGTCTCTCTGGTTGAAAAGGGGAGCTG	2460
Db	2401	GTGTAAACCCCAATATAGAAACTAGATTAAAGGGAGTCTCTCTGGTTGAAAAGGGGAGCTG	2460
Qy	2461	AGTACCTCTGGAACTGGAGCCACTCTGAAAAAAGCAAACTGAAAAACCAAGTGCCTGGG	2520
Db	2461	AGTACCTCTGGAACTGGAGCCACTCTGAAAAAAGCAAACTGAAAAACCAAGTGCCTGGG	2520
Qy	2521	TCACGTGTTACTTCCTATAAGACAGTTTAAAGTGAGACCTCGAAAAAACAATTTGTCTTACCTT	2580
Db	2521	TCACGTGTTACTTCCTATAAGACAGTTTAAAGTGAGACCTCGAAAAAACAATTTGTCTTACCTT	2580
Qy	2581	GAATAGATAGGTTTTTATGTTTGGTATATAAGAAATAAAACTTAACCTTAACCCCTGAGAC	2640
Db	2581	GAATAGATAGGTTTTTATGTTTGGTATATAAGAAATAAAACTTAACCTTAACCCCTGAGAC	2640
Qy	2641	TTTACAGGTGTGTTTATTCATATGATAGTGCATATAAAAAATTTCCCTTTAGACATCAATTTTA	2700
Db	2641	TTTACAGGTGTGTTTATTCATATGATAGTGCATATAAAAAATTTCCCTTTAGACATCAATTTTA	2700
Qy	2701	GGTAAAAATAATTTGATTAGAAAAATTTGGCCAGGTGCAGCAGCTCACACCTGCAATCC	2760
Db	2701	GGTAAAAATAATTTGATTAGAAAAATTTGGCCAGGTGCAGCAGCTCACACCTGCAATCC	2760
Qy	2761	CAGGACTTTGGGAGCCGAGCGGGTGGATCACTGTAGGTCAGGGGTTCAAGACCCAGCCT	2820
Db	2761	CAGGACTTTGGGAGCCGAGCGGGTGGATCACTGTAGGTCAGGGGTTCAAGACCCAGCCT	2820
Qy	2821	G 2821	
Db	2821	G 2821	

RESULT 2
US-08-892-695-6

US-08-892-893-8
; Sequence 6, Application US/08892695A
; Patent No. 6808878

; PATENT NO. 6806678
; GENERAL INFORMATION:

APPLICANT: Gray, Joe W

APPLICANT: GRAY, JOE W

APPLICANT: COLLINS, COLLIN
APPLICANT: Hwang. Soo In

; APPLICANT: Hwang, soo il
; APPLICANT: Godfrey, Tony

APPLICANT: Godfrey, Tony
APPLICANT: Kowal, David

APPLICANT: ROWEL, David
APPLICANT: ROMMENS, Johanna

APPLICANT: Rommens, Johanna

TITLE OF INVENTION: GENES FROM THE 20013 AMPLICON AND THEIR USES

; TITLE OF INVENTION: GENES F
; FILE REFERENCE: 2500 124US3

```

; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: IIS/08/882 695A

```

; CURRENT APPLICATION NUMBER: US/
CURRENT FILING DATE: 1997-07-15

; CURRENT FILING DATE: 1997-07-15

; EARLIER APPLICATION NUMBER: 08/785,532

; EARLIER FILING DATE: 1997-01-17

; EARLIER APPLICATION NUMBER: 08/731,499

; EARLIER FILING DATE: 1996-10-16

; EARLIER APPLICATION NUMBER: 08/680,395

```

; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 2821
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:GCAP
US-08-892-695-6

```

Query Match	100.0%	Score 2821;	DB 4;	Length 2821;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 2821;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATCCTAAGACGACAGCTGTGGGAAGCAGACACTCGGGGAAGTGGTCTCAGGAGATGTGGGT	60	
Db	1	ATCCTAAGACGACAGCTGTGGGAAGCAGACACTCGGGGAAGTGGTCTCAGGAGATGTGGGT	60	
Qy	61	CACCTGGGGTGAAGGTGAGCTTTTCAGGGTCTCCGTCAATGCAGCTGAGTTTCTTTGGC	120	
Db	61	CACCTGGGGTGAAGGTGAGCTTTTCAGGGTCTCCGTCAATGCAGCTGAGTTTCTTTGGC	120	
Qy	121	AGGGAAATTTACCACTGAGAAAGCCCTCCGGCGAGAGCTACAACCTCAGCAAGGCCAGC	180	
Db	121	AGGGAAATTTACCACTGAGAAAGCCCTCCGGCGAGAGCTACAACCTCAGCAAGGCCAGC	180	
Qy	181	TGCTCACACCCGAGAGGTCTGTGACAGGATCTTCTCCTCGTGGATGAGAAATGAGATG	240	
Db	181	TGCTCACACCCGAGAGGTCTGTGACAGGATCTTCTCCTCGTGGATGAGAAATGAGATG	240	
Qy	241	GTAAGAGGGGAGAGATGGGAGAGTGTGTCTCACTCTGCACTCATCGCCACTTTCTGGCC	300	
Db	241	GTAAGAGGGGAGAGATGGGAGAGTGTGTCTCACTCTGCACTCATCGCCACTTTCTGGCC	300	
Qy	301	GCAGTCTTTGGGCAAGGCCCTCCACCTTCCAAACCTCGGGGTCTCATCTGTGAGAAAGC	360	
Db	301	GCAGTCTTTGGGCAAGGCCCTCCACCTTCCAAACCTCGGGGTCTCATCTGTGAGAAAGC	360	
Qy	361	TGTGGAGAAGATGTCAAGAACTAAACAAAGGACCTCATGAGCACGTGTTGTAGGAGTCAAC	420	
Db	361	TGTGGAGAAGATGTCAAGAACTAAACAAAGGACCTCATGAGCACGTGTTGTAGGAGTCAAC	420	
Qy	421	TAAAGTCTTACAGAGTTGCTGATGAGGCCAGGCAGCAGAGATAGAAAGATAGGAAC	480	
Db	421	TAAAGTCTTACAGAGTTGCTGATGAGGCCAGGCAGCAGAGATAGAAAGATAGGAAC	480	
Qy	481	TTTGGAGTCAGGCAGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAAATTCCTCATC	540	
Db	481	TTTGGAGTCAGGCAGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAAATTCCTCATC	540	
Qy	541	TGGAAATGGGGATTAATAATAGTCGTTGAGAGGAATGAATAGGATAATGTGTGTTAAGAGC	600	
Db	541	TGGAAATGGGGATTAATAATAGTCGTTGAGAGGAATGAATAGGATAATGTGTGTTAAGAGC	600	
Qy	601	AGGCATAGGGTAGACCTCCATTACAGGCTGTCTGGGCTTTCCTCCCTGTAGCCCAAGGCC	660	
Db	601	AGGCATAGGGTAGACCTCCATTACAGGCTGTCTGGGCTTTCCTCCCTGTAGCCCAAGGCC	660	
Qy	661	AGCCTCAGGGGTATGTGGGAGAGAGCTGGCTTGGAATACACACTTGAGCCCTCCAGCTC	720	
Db	661	AGCCTCAGGGGTATGTGGGAGAGAGCTGGCTTGGAATACACACTTGAGCCCTCCAGCTC	720	
Qy	721	TCTCAGCTCCACCCAGCAATTTCCGTGGTATCATGGCGAAAAGTAAACCTCAATTCATCA	780	
Db	721	TCTCAGCTCCACCCAGCAATTTCCGTGGTATCATGGCGAAAAGTAAACCTCAATTCATCA	780	
Qy	781	GCAAGAAAGCCCTTAAAGGTGCGAGAGACTCTGGAGATTTCAGACACCTGCACAGCC	840	
Db	781	GCAAGAAAGCCCTTAAAGGTGCGAGAGACTCTGGAGATTTCAGACACCTGCACAGCC	840	
Qy	841	GCAAGCTTTGAGGTCAGACTGCAAGGATAGTTGGCATAGACGTGTGAGCGCATCTCTGGG	900	


```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174743
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-174743

Query Match
Best Local Similarity 9.6%; Score 270.8; DB 4; Length 601;
Matches 299; Conservative 1; Mismatches 28; Indels 2; Gaps 1;

QY 1509 AGACTGAAGAAAGCAGGAGTTGGCGGCACCGTGGCTCAGCCTGTAAATCCAGCACT 1568
DB 465 ATAATAATAATAAAAAAACAAGCGCGCGCGTGGCTCAGCCTGTAAATCCAGCACT 406
QY 1569 TTGGGAGGCCGAGCGCGGCGAGATCAGAGGTGAGGATCGAGACCATCTCGCTAAACAC 1628
DB 405 TTGGGAGGCCGAGCGCGGCGGATCAGAGGTGAGGATCAAGACCATCTCGCTAAACAC 346
QY 1629 GGGGTGAACCCCGTCTCTACTAAAAATCAAAAAATCAGCGGGTGAGGTGGCGGCGC 1688
DB 345 --GGTGAACCCCGTCTCTACTAAAAATCAAAAAATAGCGGCGAGGTGGCGGCGC 288
QY 1689 CTGTAGTCCAGCTACTCAGAGGCTGAGGCAAGAGATGGGTGAACCCAGCGGCGCG 1748
DB 287 CTGTAGTCCAGCTACTCAGAGGCTGAGGCAAGAGATGGGTGAACCCAGCGGCGCG 228
QY 1749 AGCTTACAGTGAAGCGAGATTCGGCCACTGCACTCAGCTGAGCAGTCAAGTCCG 1808
DB 227 AGCTTACAGTGAAGCGAGATTCGGCCACTGCACTCAGCTGAGCAGTCAAGTCCG 168
QY 1809 TCTCAAAAAAAGAAAAAGAAAGAAAGAA 1838
DB 167 TCTCAAAAAAAGAAAAAGAAAGAAACAA 138

RESULT 4
US-09-949-016-17189
; Sequence 17189, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17189
; LENGTH: 255679
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)---(255679)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17189

Query Match
Best Local Similarity 9.6%; Score 270.6; DB 4; Length 255679;
Matches 299; Conservative 1; Mismatches 28; Indels 2; Gaps 1;

QY 1509 AGACTGAAGAAAGCAGGAGTTGGCGGCACCGTGGCTCAGCCTGTAAATCCAGCACT 1568
DB 12325 ATAATAATAATAAAAAAACAAGCGCGCGGCTCAGCCTGTAAATCCAGCACT 12265
QY 1569 TTGGGAGGCCGAGCGCGGCGAGATCAGAGGTGAGGATCGAGACCATCTCGCTAAACAC 1628
DB 12265 TTGGGAGGCCGAGCGCGGCGGATCAGAGGTGAGGATCAAGACCATCTCGCTAAACAC 12206
QY 1629 GGGGTGAACCCCGTCTCTACTAAAAATCAAAAAATCAGCGGGTGAGGTGGCGGCGC 1688
DB 12205 --GGTGAACCCCGTCTCTACTAAAAATCAAAAAATAGCGGCGAGGTGGCGGCGC 12148
QY 1689 CTGTAGTCCAGCTACTCAGAGGCTGAGGCAAGAGATGGGTGAACCCAGCGGCGCG 1748
DB 12147 CTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGATGGGTGAACCCAGCGGCGCG 12088
```


QY 1515 AAGAAAGCAGGAGTTGCGCGGCGAGTGGTGGTCTACGCTGTAAATCCAGCACTTTGGGA 1574
Db 18825 AAAAAAAAAAAAAAAAAAGCGCGCGGTGGTCTACGCTGTAAATCCAGCACTTTGGGA 18766
QY 1575 GCGCAGCGCGGCGAGATCAGAGGTCAGGAGATCAGACCATCTGGCTTAACACGCGGTG 1634
Db 18765 GCGCAGCGCGGTGATCAGAGGTCAGGAGATCAGACCATCTGGCTTAACAC--GGTG 18708
QY 1635 AAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGTGAGGTGGCGGCGCTGTAG 1694
Db 18707 AAACCCCGTCTCTACTAAAAATACAAAAATAGCGCGGAGGTGGCGGCGCTGTAG 18648
QY 1695 TCCAGCTACTCAGAGGCTGAGGCAAGAGATGGGTGAACCCAGGCGGCGGAGCCTA 1754
Db 18647 TCCTAGCTACTCGGAGGTCAGGCGAGGAGATGGCGTGAACCCAGGAGGCGGAGCCTG 18588
QY 1755 CAGTCAGCGGAGATGGCGCACTGCACCTCCAGCTGGAGCAGAGTCCGCTCAAA 1814
Db 18587 CAGTCAGCGGAGATGGCGCACTGCACCTCCAGCTGGAGCAGAGTCCGCTCAAA 18528
QY 1815 AAAAAAAAAAGAAAGAAAGAAAGGCGAGT 1848
Db 18527 AAAAAAAAAAGAAAGAAATTTGGACAGATG 18494

RESULT 9
US-09-949-016-12234/c
; Sequence 12234, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12234
; LENGTH: 91665
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12234

Query Match 9.5%; Score 267.2; DB 4; Length 91665;
Best Local Similarity 89.5%; Pred. No. 2.3e-67;
Matches 299; Conservative 0; Mismatches 33; Indels 2; Gaps 1;
QY 1515 AAGAAAGCAGGAGTTGCGCGGCGAGTGGTGGTCTACGCTGTAAATCCAGCACTTTGGGA 1574
Db 38324 AAAAAAAAAAAAAAAAAAGCGCGCGGTGGTCTACGCTGTAAATCCAGCACTTTGGGA 38265
QY 1575 GCGCAGCGCGGCGAGATCAGAGGTCAGGAGATCAGACCATCTGGCTTAACACGCGGTG 1634
Db 38264 GCGCAGCGCGGTGATCAGAGGTCAGGAGATCAGACCATCTGGCTTAACAC--GGTG 38207
QY 1635 AAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGTGAGGTGGCGGCGCTGTAG 1694
Db 38206 AAACCCCGTCTCTACTAAAAATACAAAAATAGCGCGGAGGTGGCGGCGCTGTAG 38147
QY 1695 TCCAGCTACTCAGAGGCTGAGGCAAGAGATGGGTGAACCCAGGCGGCGGAGCCTA 1754
Db 38146 TCCTAGCTACTCGGAGGTCAGGCGAGGAGATGGCGTGAACCCAGGAGGCGGAGCCTG 38087
QY 1755 CAGTCAGCGGAGATGGCGCACTGCACCTCCAGCTGGAGCAGAGTCCGCTCAAA 1814
Db 38086 CAGTCAGCGGAGATGGCGCACTGCACCTCCAGCTGGGCGAGCAGGAGTCCGCTCAAA 38027

QY 1815 AAAAAAAAAAGAAAGAAAGAAAGGCGAGT 1848
Db 38026 AAAAAAAAAAGAAAGAAATTTGGACAGATG 37993
RESULT 10
US-09-949-016-15829/c
; Sequence 15829, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15829
; LENGTH: 44166
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15829

Query Match 9.5%; Score 266.6; DB 4; Length 44166;
Best Local Similarity 90.5%; Pred. No. 2.1e-67;
Matches 296; Conservative 0; Mismatches 29; Indels 2; Gaps 1;
QY 1507 TGAGACTCAAGAAAGGCGAGGTTGCGGCGGCGAGTGGTGGTCTACGCTGTAAATCCAGCA 1566
Db 25428 TCACACAAAGAAATCTCATATTTTGGCGGCGCGGTGGTCTACGCTGTAAATCCAGCA 25369
QY 1567 CTTTGGGAGCGCGGCGGCGAGATCAGAGGTCAAGAGATCGAGACCATCTGGCTTAAC 1626
Db 25368 CTTTGGGAGCGCGGCGGCGGTGGATCAGAGGTCAAGAGATCGAGACCATCTGGCTTAAC 25309
QY 1627 ACGGGGTGAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGTGAGTGGCGGCG 1686
Db 25308 AC--GGTGAACCCCGTCTCTACTAAAAATACAAAAATAGCGCGGCGAGTGGCGGCG 25251
QY 1687 GCCTGTAGTCCAGCTACTCAGGAGGCTGAGGCGAGAGATGGCTGAACCCAGGGGCG 1746
Db 25250 GCCTGTAGTCCAGCTACTCAGGAGGCTGAGGCGAGAGATGGTGTGAACCCAGGGGCG 25191
QY 1747 CGAGCTACAGTGAGCGGAGATTGGCCACTGCACTCCAGCTGGACGACAGTGGAGACTC 1806
Db 25190 AGAGCTCAGTGAGCGGAGATTGGCCACTGCACTCCAGCTGGCGGAGAGACTC 25131
QY 1807 CGTCTCAAAAAAAGAAAGAAAGAA 1833
Db 25130 CGTCTCAAAAAAAGAAAGAAAGAA 25104

RESULT 11
US-09-949-016-14894/c
; Sequence 14894, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2005, 01:06:08 ; Search time 1103.5 Seconds
(without alignments)
15978.923 Million cell updates/sec

Title: US-08-731-499-6

Perfect score: 2821

Sequence: 1 ATCTTAGACGACAGCCTG.....AGGGTTTCAGACGACGCTG 2821

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgm2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgm2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgm2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgm2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgm2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgm2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgm2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
9: /cgm2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgm2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgm2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgm2_6/ptodata/1/pubpna/US09D_PUBCOMB.seq.*
13: /cgm2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgm2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgm2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgm2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgm2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgm2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgm2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgm2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgm2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgm2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq.*
23: /cgm2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgm2_6/ptodata/1/pubpna/US11B_PUBCOMB.seq.*
25: /cgm2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgm2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2821	100.0	2821	28	US-08-731-499-6
2	1103.4	39.1	11006	22	Sequence 6, Appl
3	1103.4	39.1	11006	22	Sequence 11, Appl
4	559.6	19.8	600	13	Sequence 11, Appl
5	559.6	19.8	600	17	Sequence 134687
6	337	11.9	1506	20	Sequence 134687
7	337	11.9	1506	20	Sequence 23318, A
					Sequence 29199, A

8	277.2	9.8	63520	20	US-10-266-103-39	Sequence 39, Appl
9	275	9.7	32152	9	US-09-764-855-328	Sequence 328, App
10	275	9.7	32152	9	US-09-764-872-518	Sequence 518, App
11	275	9.7	32152	14	US-10-072-349-328	Sequence 328, App
12	275	9.7	129042	13	US-10-087-192-1240	Sequence 1240, Ap
13	274.6	9.7	335	17	US-10-242-355-840	Sequence 840, App
14	273.4	9.7	238484	13	US-10-087-192-544	Sequence 544, App
15	272.6	9.7	7995	16	US-10-220-510-5	Sequence 5, Appl
16	272.6	9.7	74279	22	US-10-737-082-16	Sequence 16, Appl
17	272.6	9.7	74279	22	US-10-765-790-16	Sequence 16, Appl
18	271.8	9.6	45000	15	US-10-000-213-12	Sequence 12, Appl
19	271.8	9.6	122186	10	US-09-563-728A-36	Sequence 36, Appl
20	271.6	9.6	24264	15	US-10-017-161-565	Sequence 665, App
21	270.8	9.6	59942	19	US-10-741-601-5775	Sequence 5775, Ap
22	270	9.6	810	10	US-09-764-891-5798	Sequence 5798, Ap
23	269.2	9.5	370469	13	US-10-087-192-250	Sequence 250, App
24	269	9.5	1537	13	US-10-027-632-261000	Sequence 261000,
25	269	9.5	1537	17	US-10-027-632-261000	Sequence 1, Appli
26	266.4	9.4	334462	21	US-10-496-011-1	Sequence 17630, A
27	266	9.4	62555	21	US-10-741-600-17630	Sequence 195045,
28	264	9.4	471	13	US-10-027-632-195045	Sequence 195045,
29	264	9.4	471	13	US-10-027-632-195046	Sequence 195046,
30	264	9.4	471	17	US-10-027-632-195045	Sequence 195045,
31	264	9.4	471	17	US-10-027-632-195046	Sequence 195046,
32	263.6	9.3	816	10	US-09-764-891-5797	Sequence 5797, Ap
33	263.6	9.3	816	10	US-09-764-891-5799	Sequence 5799, Ap
34	263.6	9.3	101569	19	US-10-717-597-10	Sequence 10, Appl
35	259.8	9.2	68495	19	US-10-322-281-750	Sequence 750, App
36	259	9.2	405660	19	US-10-322-696-82	Sequence 82, Appl
37	258.6	9.2	31304	10	US-09-764-891-9556	Sequence 9556, Ap
38	258.6	9.2	31304	15	US-10-205-428-884	Sequence 884, App
39	258.6	9.2	44030	11	US-09-997-722-172	Sequence 172, App
40	258.6	9.2	44030	16	US-10-105-612-4	Sequence 4, Appli
41	258	9.1	1263	13	US-10-027-632-258046	Sequence 258046,
42	258	9.1	1263	17	US-10-027-632-258046	Sequence 258046,
43	257.4	9.1	15306	10	US-10-027-632-10219	Sequence 10219, A
44	257	9.1	309	9	US-09-764-877-2197	Sequence 2197, Ap
45	257	9.1	309	17	US-10-242-515-2197	Sequence 2197, Ap

ALIGNMENTS

RESULT 1

US-08-731-499-6
; Sequence 6, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION: Joe W.
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOMBEL, David
; APPLICANT: ROMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/680,395
;; FILING DATE: 15-JUL-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hunter, Tom
;; REGISTRATION NUMBER: 38,498
;; REFERENCE/DOCKET NUMBER: 23070-068910
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2821 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: 1..2821 /note= "cDNA clone GCAP encodes a
;; OTHER INFORMATION: guanino cyclase activating protein
;; OTHER INFORMATION:
US-08-731-499-6

Query Match 100.0%; Score 2821; DB 8; Length 2821;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2821; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATCCTAAGACGACACAGCCTGGGAAGCCAGACACTGGGAAGTGGTCTCAGGGGATGTGGGT	60
Db	1	ATCCTAAGACGACACAGCCTGGGAAGCCAGACACTGGGAAGTGGTCTCAGGGGATGTGGGT	60
Qy	61	CACCTGGGGTGAAGTGGAGCTTTCAGGGTCTCCCGTCAATGAGCTGAGTTTCTTTGGC	120
Db	61	CACCTGGGGTGAAGTGGAGCTTTCAGGGTCTCCCGTCAATGAGCTGAGTTTCTTTGGC	120
Qy	121	AGGGAATTTACAGCTGAAGAAAGCCTCCGCGGAGAGCTACAACTCAGCAAGCCAGC	180
Db	121	AGGGAATTTACAGCTGAAGAAAGCCTCCGCGGAGAGCTACAACTCAGCAAGCCAGC	180
Qy	181	TGCTCACACCCGAGAGTCTGGACAGGATCTTCTCCTGGTGGATGAGAATGGAGATG	240
Db	181	TGCTCACACCCGAGAGTCTGGACAGGATCTTCTCCTGGTGGATGAGAATGGAGATG	240
Qy	241	GTAAGAGGGGAGAGATGGGGAGAGTGTCTCCACTCTGCATCATCGCCACTTTCTGGCC	300
Db	241	GTAAGAGGGGAGAGATGGGGAGAGTGTCTCCACTCTGCATCATCGCCACTTTCTGGCC	300
Qy	301	GCAGTCTCTGGCAAGGCCCTCCACCTTCCAACTGGGGTCCCTCATCTGTGAGAGGC	360
Db	301	GCAGTCTCTGGCAAGGCCCTCCACCTTCCAACTGGGGTCCCTCATCTGTGAGAGGC	360
Qy	361	TGTGAGAAGATGTCAATGAACCTAAAGAGGACTCATGAGCAGCGTTTGTAGGAGTGAC	420
Db	361	TGTGAGAAGATGTCAATGAACCTAAAGAGGACTCATGAGCAGCGTTTGTAGGAGTGAC	420
Qy	421	TAAAGTCTCAGAGGTTGCTGATGGAGGCCAGGCAGCGAATAGAAAGATAGGAAC	480
Db	421	TAAAGTCTCAGAGGTTGCTGATGGAGGCCAGGCAGCGAATAGAAAGATAGGAAC	480
Qy	481	TTTGGAGTCAGCGAGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAATTTCTCATC	540
Db	481	TTTGGAGTCAGCGAGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAATTTCTCATC	540
Qy	541	TGGAATTTGGGGATAAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAGAGC	600
Db	541	TGGAATTTGGGGATAAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAGAGC	600
Qy	601	AGGCATAGGGTAGACCTCCATTGAGGCTGCTTGGGCTTTCTCCCTGTAGGCCCAAGGCC	660
Db	601	AGGCATAGGGTAGACCTCCATTGAGGCTGCTTGGGCTTTCTCCCTGTAGGCCCAAGGCC	660

Qy	661	AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGAAATACACACTTGAGCCCTCCAGCTC	720
Db	661	AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGAAATACACACTTGAGCCCTCCAGCTC	720
Qy	721	TCTCAGCTCCACCAGCATTTTCCGTGTPACCATCGCAAAAGTAAATCTCAATTCATCA	780
Db	721	TCTCAGCTCCACCAGCATTTTCCGTGTPACCATCGCAAAAGTAAATCTCAATTCATCA	780
Qy	781	GCAAGAAAGCCCTTAAAGTGCAGAGAGACTCTGGAGATTGAGACACCTGCAAGCC	840
Db	781	GCAAGAAAGCCCTTAAAGTGCAGAGAGACTCTGGAGATTGAGACACCTGCAAGCC	840
Qy	841	GCAAGCTTGAGTCTGAGACTGCAAGGATAGTTGGCATAAAGCTGTAGCGCATCCTGGG	900
Db	841	GCAAGCTTGAGTCTGAGACTGCAAGGATAGTTGGCATAAAGCTGTAGCGCATCCTGGG	900
Qy	901	AGCAGGCTCTCTCTCTGCCCCCAGACCAGGCTCTCCCTTCTTATACATGACCACTC	960
Db	901	AGCAGGCTCTCTCTCTGCCCCCAGACCAGGCTCTCCCTTCTTATACATGACCACTC	960
Qy	961	TCCTCCCCCTTGCTCAGCCAGCTGTCTCTGAAAGGATTTGTAAGGTGCGCTCGGA	1020
Db	961	TCCTCCCCCTTGCTCAGCCAGCTGTCTCTGAAAGGATTTGTAAGGTGCGCTCGGA	1020
Qy	1021	CAAGTGGGTGATGAAGATGCTGCAGATGACATGAATCCAGCAGCTGCTCGCTCAGCA	1080
Db	1021	CAAGTGGGTGATGAAGATGCTGCAGATGACATGAATCCAGCAGCTGCTCGCTCAGCA	1080
Qy	1081	GAGACGGAAAAAGTCCATGTTCTGAGGAGTCTGGGGCCCCCTCCACGACTCCAGGCTCACC	1140
Db	1081	GAGACGGAAAAAGTCCATGTTCTGAGGAGTCTGGGGCCCCCTCCACGACTCCAGGCTCACC	1140
Qy	1141	CAGGTTTCCAGGCTAGTAGAGGCTCCCTGGCTCAGCTCAGCTCATGCCCACTCTTCCCC	1200
Db	1141	CAGGTTTCCAGGCTAGTAGAGGCTCCCTGGCTCAGCTCAGCTCATGCCCACTCTTCCCC	1200
Qy	1201	TGGTGTGACTTCTTGGCACCCTGTGAGGGCTGAGTGGGGATGGGGAAGGGCTGTG	1260
Db	1201	TGGTGTGACTTCTTGGCACCCTGTGAGGGCTGAGTGGGGATGGGGAAGGGCTGTG	1260
Qy	1261	GTTTGAAGTGGCCAAAGGGCATAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGG	1320
Db	1261	GTTTGAAGTGGCCAAAGGGCATAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGG	1320
Qy	1321	AATTCACTTACTTTTCTGTTTCCAGCTCTGAGGAGCTGCTGCTTGGCTGGGTGGT	1380
Db	1321	AATTCACTTACTTTTCTGTTTCCAGCTCTGAGGAGCTGCTGCTTGGCTGGGTGGT	1380
Qy	1381	GTGGGGCTCCCACTGTTTCTGGGTGTTCTCAAGTGGGAAGCAAGAGCCAACTGAGGGGTGA	1440
Db	1381	GTGGGGCTCCCACTGTTTCTGGGTGTTCTCAAGTGGGAAGCAAGAGCCAACTGAGGGGTGA	1440
Qy	1441	GGGTCCCAAGCAAAATCAGAAATGAGAACACAAAGCTGGTAGGAGGCAAGGGGTGGGA	1500
Db	1441	GGGTCCCAAGCAAAATCAGAAATGAGAACACAAAGCTGGTAGGAGGCAAGGGGTGGGA	1500
Qy	1501	GGGTGTTGAGACTGAAGAAAGGAGGTTGCCGGGACCGGTGGCTCACCGCTGTAATC	1560
Db	1501	GGGTGTTGAGACTGAAGAAAGGAGGTTGCCGGGACCGGTGGCTCACCGCTGTAATC	1560
Qy	1561	CCAGCACTTTGGGAGGCCGAGCGGGCAGATCAGAGGTGAGGAGATCGAGACCATCTCTG	1620
Db	1561	CCAGCACTTTGGGAGGCCGAGCGGGCAGATCAGAGGTGAGGAGATCGAGACCATCTCTG	1620
Qy	1621	GCTAACACGGGGTGAACCCCGTCTCTACTAAATAACAAAAATCAGCCCGGGTGAAGTG	1680
Db	1621	GCTAACACGGGGTGAACCCCGTCTCTACTAAATAACAAAAATCAGCCCGGGTGAAGTG	1680
Qy	1681	GCGGCGCTCTAGTCCAGCTACTCAGGAGGCTGAGCAAGAGAAATGCGCTGCAACCCCA	1740
Db	1681	GCGGCGCTCTAGTCCAGCTACTCAGGAGGCTGAGCAAGAGAAATGCGCTGCAACCCCA	1740
Qy	1741	GGGGGCGGAGCTTACAGTGAAGCGAGATTGCGCCACTGCACTCCAGCCTGGACGAGTG	1800

Db 1741 GGGGGCCGAGCCTACAGTGAGCGAGATTGCGCACTGCACTCCAGCCTGGACGACAGTG 1800
Qy 1801 AGACTCCGCTCTCAAAAAAAGAAAAAGAAAAAGGAGGAGGAGTGTGGGGGGCAG 1860
Db 1801 AGACTCCGCTCTCAAAAAAAGAAAAAGAAAAAGGAGGAGGAGTGTGGGGGGCAG 1860
Qy 1861 GGGGAGCAATTAATCTTATAAATTCCTGGGATGCTGAGGGCGGTTCATGGGAGGACCTTG 1920
Db 1861 GGGGAGCAATTAATCTTATAAATTCCTGGGATGCTGAGGGCGGTTCATGGGAGGACCTTG 1920
Qy 1921 GCCTCCTCTCCCAAGGATCCTCACAGTGTGTCAACAGGAAAAATGGCAGCAATA 1980
Db 1921 GCCTCCTCTCCCAAGGATCCTCACAGTGTGTCAACAGGAAAAATGGCAGCAATA 1980
Qy 1981 CGCTGAGGCTGTGGTCTTCTGCTTTGAAAGGGTCAGCTGTACTTAAAGGACTGTGTT 2040
Db 1981 CGCTGAGGCTGTGGTCTTCTGCTTTGAAAGGGTCAGCTGTACTTAAAGGACTGTGTT 2040
Qy 2041 CAGCTGTGCTGGTGTCTGCTGGGACCCCTGCTGCAACCCACCACTCTCCCCCAACAA 2100
Db 2041 CAGCTGTGCTGGTGTCTGCTGGGACCCCTGCTGCAACCCACCACTCTCCCCCAACAA 2100
Qy 2101 TCCTCTCTTCCATCCATATCCCAAGTATGGACCTTCCACAACTCCCAAGCCATAAGCTG 2160
Db 2101 TCCTCTCTTCCATCCATATCCCAAGTATGGACCTTCCACAACTCCCAAGCCATAAGCTG 2160
Qy 2161 AATGTTTCTCTTAAAGATGGAGAAAACTTCTGCTGTCTCTGGCAAGAAATTTGGGGAC 2220
Db 2161 AATGTTTCTCTTAAAGATGGAGAAAACTTCTGCTGTCTCTGGCAAGAAATTTGGGGAC 2220
Qy 2221 TGTGACTGGGATGTGGGCTGGCTGCTTAACTGCTGTGTGACCCCAAGACAGCCA 2280
Db 2221 TGTGACTGGGATGTGGGCTGGCTGCTTAACTGCTGTGTGACCCCAAGACAGCCA 2280
Qy 2281 CTTCTCTCCCTAACTGTTATGCTTGGCAGCAGTGGAGCTGGGCTGGGCTAGGCGA 2340
Db 2281 CTTCTCTCCCTAACTGTTATGCTTGGCAGCAGTGGAGCTGGGCTGGGCTAGGCGA 2340
Qy 2341 ACAGTTTGTGATATGCTGTTTATGATGCTGAAATTAATTTTGTATATAAATCTTAT 2400
Db 2341 ACAGTTTGTGATATGCTGTTTATGATGCTGAAATTAATTTTGTATATAAATCTTAT 2400
Qy 2401 GTGTAACCCCAATATAGAAACTAGATTAAGAGGAGTCTCTCTGTTGAAAGGGGAGCTG 2460
Db 2401 GTGTAACCCCAATATAGAAACTAGATTAAGAGGAGTCTCTCTGTTGAAAGGGGAGCTG 2460
Qy 2461 AGTACCTCTGAACTGGAGGACCTCTGAAAGAGCAAACTGAAACCAAGTGCCTGGG 2520
Db 2461 AGTACCTCTGAACTGGAGGACCTCTGAAAGAGCAAACTGAAACCAAGTGCCTGGG 2520
Qy 2521 TCACCTGTTACTCTATAAGACAGTTTAAAGTGAGACCTGGAATAAATTTGCTTTACCTT 2580
Db 2521 TCACCTGTTACTCTATAAGACAGTTTAAAGTGAGACCTGGAATAAATTTGCTTTACCTT 2580
Qy 2581 GAATAGATAGGTTTATGTTGATATAGAAATAAATAACCTATTAAACCTTGAGAC 2640
Db 2581 GAATAGATAGGTTTATGTTGATATAGAAATAAATAACCTATTAAACCTTGAGAC 2640
Qy 2641 TTTACAGGTGTGTTATTTATATAGTAGTATATAAATAATTTCTTTAGACATCAATTTTA 2700
Db 2641 TTTACAGGTGTGTTATTTATATAGTAGTATATAAATAATTTCTTTAGACATCAATTTTA 2700
Qy 2701 GGTAAAAAATAATGATAGAAAAATTTGGCCAGGTGAGCAGCTCACACCTGCAATCC 2760
Db 2701 GGTAAAAAATAATGATAGAAAAATTTGGCCAGGTGAGCAGCTCACACCTGCAATCC 2760
Qy 2761 CAGGACTTTGGAGGCCGAGGGGTGATCACTGAGGTCAAGGCTTCAAGACCAAGCT 2820
Db 2761 CAGGACTTTGGAGGCCGAGGGGTGATCACTGAGGTCAAGGCTTCAAGACCAAGCT 2820
Qy 2821 G 2821

Db 2821 G 2821
RESULT 2
US-10-737-082-11
; Sequence 11, Application US/10737082
; Publication No. US20050130170A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2032
; CURRENT APPLICATION NUMBER: US/10/737,082
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 11006
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-737-082-11

Query Match 39.1%; Score 1103.4; DB 22; Length 11006;
Best Local Similarity 99.9%; Pred. No. 1.4e-299;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATCTTAAGACGACAGCCTGGGAAGCCAGCAGCTGGGAAAGTGTGCTGAGGAGTGGGT 60
Db 9902 ATCTTAAGACGACAGCCTGGGAAGCCAGCAGCTGGGAAAGTGTGCTGAGGAGTGGGT 9961
Qy 61 CACTGGGTGAAGGTGGAGCTTTCAGGGTCTCCCGTCAATGACAGTGTGCTTTTGGC 120
Db 9962 CACTGGGTGAAGGTGGAGCTTTCAGGGTCTCCCGTCAATGACAGTGTGCTTTTGGC 10021
Qy 121 AGGGAATTTACAGCTGAAGAAAGCTCCGCGGAGAGCTACAACTGAGCAAGGCCAGC 180
Db 10022 AGGGAATTTACAGCTGAAGAAAGCTCCGCGGAGAGCTACAACTGAGCAAGGCCAGC 10081
Qy 181 TGTCAACCCCGAGGAGTGTGGACAGGATCTTCTCTGTTGGATGAGAAATGGAGATG 240
Db 10082 TGTCAACCCCGAGGAGTGTGGACAGGATCTTCTCTGTTGGATGAGAAATGGAGATG 10141
Qy 241 GTAAGAGGGCAGAGATGGGAGAGTGTCTCCACTCTGCATCATGCCACTTTCTGGCC 300
Db 10142 GTAAGAGGGCAGAGATGGGAGAGTGTCTCCACTCTGCATCATGCCACTTTCTGGCC 10201
Qy 301 GCAGTCTCTGGGCAAGGCCCTCCACCTTCCAACTGGGGTCTCTCTGTGAGAGGC 360
Db 10202 GCAGTCTCTGGGCAAGGCCCTCCACCTTCCAACTGGGGTCTCTCTGTGAGAGGC 10261
Qy 361 TGTGGAAGAGTGTCAATGAATTAACAAAGGAGCTCATGAGCAGCTGTTGTAGAGTGAC 420
Db 10262 TGTGGAAGAGTGTCAATGAATTAACAAAGGAGCTCATGAGCAGCTGTTGTAGAGTGAC 10321
Qy 421 TAAAGTCTTACAGGAGTGTGATGGAGGCCAGGCCAGCAATAGAAATAGGAAC 480
Db 10322 TAAAGTCTTACAGGAGTGTGATGGAGGCCAGGCCAGCAATAGAAATAGGAAC 10381
Qy 481 TTTGGAGTCAGGAGGAGTGTATATTGAGCTTCTCGTCTTAGTCTCAATTTCTTCATC 540
Db 10382 TTTGGAGTCAGGAGGAGTGTATATTGAGCTTCTCGTCTTAGTCTCAATTTCTTCATC 10441
Qy 541 TGGAAAAATGGGATTAATAATAGTGGTTGAGAGGAATGAATAGGATTAATGTGTTAAGAGC 600
Db 10442 TGGAAAAATGGGATTAATAATAGTGGTTGAGAGGAATGAATAGGATTAATGTGTTAAGAGC 10501


```

QY 601 AGGCATAGGGTAGA CTTCCATTCAGGCTGCTTGGGCTTTCCTCCCTGAGCCCAAGGCC 660
DB 10502 AGGCATAGGGTAGA CTTCCATTCAGGCTGCTTGGGCTTTCCTCCCTGAGCCCAAGGCC 10561
QY 661 AGCCTCAGGGCTATGTGGGGAGAGAGCTGGCTTGGGAATACACACTTGAAGCCCTCAGGCTC 720
DB 10562 AGCCTCAGGGCTATGTGGGGAGAGAGCTGGCTTGGGAATACACACTTGAAGCCCTCAGGCTC 10621
QY 721 TCTCAGCTCCACCCAGCATTTCCTGGGTACCATGCGCAAAAGTAAACTTCAATTCATCA 780
DB 10622 TCTCAGCTCCACCCAGCATTTCCTGGGTACCATGCGCAAAAGTAAACTTCAATTCATCA 10681
QY 781 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTCTGGAGATTCAAGACCTGACAAAGCC 840
DB 10682 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTCTGGAGATTCAAGACCTGACAAAGCC 10741
QY 841 GCAAGCTTGAGGTCTGAGATCAGCATGAGATGATGAGCATAAAGCTGTAGGCGCATCCTGGG 900
DB 10742 GCAAGCTTGAGGTCTGAGATCAGCATGAGATGATGAGCATAAAGCTGTAGGCGCATCCTGGG 10801
QY 901 AGCGAGGTCTCTCCTCCCTGCCCCCAGACCCAGGCTCTCCCTTCTTCTACATGACCACTTC 960
DB 10802 AGCGAGGTCTCTCCTCCCTGCCCCCAGACCCAGGCTCTCCCTTCTTCTACATGACCACTTC 10861
QY 961 TCCTCCCTCTGCTCAGGCCAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCGCTCGGGA 1020
DB 10862 TCCTCCCTCTGCTCAGGCCAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCGCTCGGGA 10921
QY 1021 CAAAGTGGGTGATGAAGATGCTCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCA 1080
DB 10922 CAAAGTGGGTGATGAAGATGCTCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCA 10981
QY 1081 GAGACGGAAGAGTGCATGTTCTGA 1105
DB 10982 GAGACGGAAGAGTGCATGTTCTGA 11006

```

RESULT 3

```

US-10-765-790-11
; Sequence 11, Application US/10765790
; Publication No. US20050130172A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John P.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US/10/765,790
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 11006
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-765-790-11

```

```

Query Match 39.1%; Score 1103.4; DB 22; Length 11006;
Best Local Similarity 99.9%; Pred. No. 1.4e-299;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCCTAAGACGACAGCCTGGGAAGCCAGCAGCTGGGGAAGTGCTGCTGAGGAGATGTTGGGT 60
DB 9902 ATCCTAAGACGACAGCCTGGGAAGCCAGCAGCTGGGGAAGTGCTGCTGAGGAGATGTTGGGT 9961
QY 61 CACTGGGGTGAAGTGGAGCTTTTCAGGGTCTCCCGTCAATGACGTGAGTTTCTTTGGC 120

```

```

DB 9962 CACTGGGGTGAAGTGGAGCTTTTCAGGGTCTCCCGTCAATGACGTGAGTTTCTTTTGGC 10021
QY 121 AGGGATTTTACAGCTGAAGAAAGCCTCCGCGAGAGCTACAACTGAGCAAGGCCAGC 180
DB 10022 AGGGATTTTACAGCTGAAGAAAGCCTCCGCGAGAGCTACAACTGAGCAAGGCCAGC 10081
QY 181 TGCTCACACCCAGAGAGTCTGTGGACAGGATCTTCTCTCTGTGGATGAGAAATGAGATG 240
DB 10082 TGCTCACACCCAGAGAGTCTGTGGACAGGATCTTCTCTCTGTGGATGAGAAATGAGATG 10141
QY 241 GTAAAGGGGACAGATGGGGAGAGTGTGTCCACTCTGTGCATCATCGCCACCTTTCTGGCC 300
DB 10142 GTAAAGGGGACAGATGGGGAGAGTGTGTGTCCACTCTGTGCATCATCGCCACCTTTCTGGCC 10201
QY 301 GCACGCTCTTGGGCAAGGCCCTCCACCTTCCAAACCTGGGGTCCCTCATCTGTGAGAGGC 360
DB 10202 GCACGCTCTTGGGCAAGGCCCTCCACCTTCCAAACCTGGGGTCCCTCATCTGTGAGAGGC 10261
QY 361 TGTGAGAGAGATGTCTCAACTTAAACAAAGGAGCTCATGAGCACGTGTTTGTAGGAGTGAC 420
DB 10262 TGTGAGAGAGATGTCTCAACTTAAACAAAGGAGCTCATGAGCACGTGTTTGTAGGAGTGAC 10321
QY 421 TAAAGTCTTACAGAGTTGTGTATGGAGGCGAGCAAGCAAGATAGAAATAGGAAC 480
DB 10322 TAAAGTCTTACAGAGTTGTGTATGGAGGCGAGCAAGCAAGATAGGAATAGGAAC 10381
QY 481 TTTGGAGTCAAGCAGGAGTGATATATTGAGCTTCTCGTCTCTAGTCTCAATTTCTCTCATC 540
DB 10382 TTTGGAGTCAAGCAGGAGTGATATATTGAGCTTCTCGTCTCTAGTCTCAATTTCTCTCATC 10441
QY 541 TGGAAATGGGGATTAATAATAGTGTTCAGAGGAATGAATAGGATAATGTGTTTAAGAGC 600
DB 10442 TGGAAATGGGGATTAATAATAGTGTTCAGAGGAATGAATAGGATAATGTGTTTAAGAGC 10501
QY 601 AGGCATAGGGTAGA CTTCCATTCAGGCTGCTTGGGCTTTCCTCCCTGAGCCCAAGGCC 660
DB 10502 AGGCATAGGGTAGA CTTCCATTCAGGCTGCTTGGGCTTTCCTCCCTGAGCCCAAGGCC 10561
QY 661 AGCCTCAGGGCTATGTGGGGAGAGAGCTGGCTTGGAAATACACACTTGAAGCCCTCAGGCTC 720
DB 10562 AGCCTCAGGGCTATGTGGGGAGAGAGCTGGCTTGGAAATACACACTTGAAGCCCTCAGGCTC 10621
QY 721 TCTCAGCTCCACCCAGCATTTCCTGGGTACCATGCGCAAAAGTAAACTTCAATTCATCA 780
DB 10622 TCTCAGCTCCACCCAGCATTTCCTGGGTACCATGCGCAAAAGTAAACTTCAATTCATCA 10681
QY 781 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTCTGGAGATTCAAGACCTGACAAAGCC 840
DB 10682 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTCTGGAGATTCAAGACCTGACAAAGCC 10741
QY 841 GCAAGCTTGAGGTCTGAGATGAGATGATGAGCATAAAGCTGTAGGCGCATCCTGGG 900
DB 10742 GCAAGCTTGAGGTCTGAGATGAGATGATGAGCATAAAGCTGTAGGCGCATCCTGGG 10801
QY 901 AGCGAGGTCTCTCCTCCCTGCCCCCAGACCCAGGCTCTCCCTTCTTCTACATGACCACTTC 960
DB 10802 AGCGAGGTCTCTCCTCCCTGCCCCCAGACCCAGGCTCTCCCTTCTTCTACATGACCACTTC 10861
QY 961 TCCTCCCTCTGCTCAGGCCAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCGCTCGGGA 1020
DB 10862 TCCTCCCTCTGCTCAGGCCAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCGCTCGGGA 10921
QY 1021 CAAAGTGGGTGATGAAGATGCTCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCA 1080
DB 10922 CAAAGTGGGTGATGAAGATGCTCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCA 10981
QY 1081 GAGACGGAAGAGTGCATGTTCTGA 1105
DB 10982 GAGACGGAAGAGTGCATGTTCTGA 11006

```

```

; Sequence 134687, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134687
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134687

Query Match          19.8%; Score 559.6; DB 13; Length 600;
Best Local Similarity 97.5%; Pred. No. 8e-147;
Matches 589; Conservative 1; Mismatches 10; Indels 4; Gaps 2;

QY 1068 GGCTGCTCAGCAGACGCGAAAGTGCATGTTCTGAGGAGTCTGGGGCCCTCCACGA 1127
DB 1 GGCTGCTCAGCAGACGCGAAAGTGCATGTTCTGAGGAGTCTGGGGCCCTCCACGA 60

QY 1128 CTCACGGCTCACCCAGGTTTCCAGGAGTGTAGAGGGTCCCTGGCTCAGGCTGCTCATG 1187
DB 61 CTCACGGCTCACCCAGGTTTCCAGGAGTGTAGAGGGT--CCTGGCTCAGGCTGCTCATG 118

QY 1188 CCCACTCTTCCCTGCTGTGTTGACTTCTCGGCACCCCTCTGAGGGTCTGAGTGGGATGG 1247
DB 119 CCCACTCTTCCCTGCTGTGTTGACTTCTCGGCACCCCTCTGAGGGTCTGAGTGGGATGG 178

QY 1248 GGAAGGGCTGCTGGGTTTGAAGTGGCCAAACAGGCGATAGTCCATTTTGAGGAGTCCCTG 1307
DB 179 GGAAGGGCTGCTGGGTTTGAAGTGGCCAAACAGGCGATAGTCCATTTTGAGGAGTCCCTG 238

QY 1308 GGATGGTGAAGGGAATTCAGTTTACTTTTCTGAGCGCTCTCTGGGAGGACTGTGCCT 1367
DB 239 GGATGGTGAAGGGAATTCAGTTTACTTTTCTGAGCGCTCTCTGGGAGGACTGTGCCT 298

QY 1368 TGCGTGGGTGGTGTGGGGTCCCAAGTTCCTGGGTGTTCTCAGTTGGAAGCAAGAGCC 1427
DB 299 TGCGTGGGTGGTGTGGGGTCCCAAGTTCCTGGGTGTTCTCAGTTGGAAGCAAGAGCC 358

QY 1428 AACTCAGGGGTGAGGGTCCCAAGTTCCTGGGTGTTCTCAGTTGGAAGCAAGAGCTGTGAG 1487
DB 359 AACTCAGGGGTGAGGGTCCCAAGTTCCTGGGTGTTCTCAGTTGGAAGCAAGAGCTGTGAG 418

QY 1488 GGCAGGGGTGGAGGGTGTGAGCTGAAGAAAGCAGAGTTCGCGGGCAGGTGGCT 1547
DB 419 GGCAGGGGTGGAGGGTGTGAGCTGAAGAAAGCAGAGTTCGCGGGCAGGTGGCT 478

QY 1548 CACGCTGTAAATCCAGCACTTTGGGAGGCGCAGGCGGCGAGATCACAGGTCAGGAGAT 1607
DB 479 CAYGCTGTAAATCCAGCACTTTAGGAGCGCGGCGGCGAGATCACAGGTCAGGAGAT 538

QY 1608 CGAGACCATCTCGGTGTAACCGGGGTGAACCCCGTCTCTACTAAATAATCAAAAAATCA 1667
DB 1608 CGAGACCATCTCGGTGTAACCGGGGTGAACCCCGTCTCTACTAAATAATCAAAAAATCA 1667

```

```

DB 539 CGAGACCATCTCGGTGTAACACCGTGTGAAA--CTAATACTAATAAATAAAAAAATAA 596
QY 1668 GCCG 1671
DB 597 GCCG 600

RESULT 5
US-10-027-632-134687
; Sequence 134687, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134687
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134687

Query Match          19.8%; Score 559.6; DB 17; Length 600;
Best Local Similarity 97.5%; Pred. No. 8e-147;
Matches 589; Conservative 1; Mismatches 10; Indels 4; Gaps 2;

QY 1068 GGCTGCTCAGCAGACGCGAAAGTGCATGTTCTGAGGAGTCTGGGGCCCTCCACGA 1127
DB 1 GGCTGCTCAGCAGACGCGAAAGTGCATGTTCTGAGGAGTCTGGGGCCCTCCACGA 60

QY 1128 CTCACGGCTCACCCAGGTTTCCAGGAGTGTAGAGGGTCCCTGGCTCAGGCTGCTCATG 1187
DB 61 CTCACGGCTCACCCAGGTTTCCAGGAGTGTAGAGGGT--CCTGGCTCAGGCTGCTCATG 118

QY 1188 CCCACTCTTCCCTGCTGTGTTGACTTCTGGGCACCCCTCTGAGGGTCTGAGTGGGATGG 1247
DB 119 CCCACTCTTCCCTGCTGTGTTGACTTCTGGGCACCCCTCTGAGGGTCTGAGTGGGATGG 178

QY 1248 GGAAGGGCTGCTGGGTTTGAAGTGGCCAAACAGGCGATAGTCCATTTTGAGGAGTCCCTG 1307
DB 179 GGAAGGGCTGCTGGGTTTGAAGTGGCCAAACAGGCGATAGTCCATTTTGAGGAGTCCCTG 238

QY 1308 GGATGGTGAAGGGAATTCAGTTTACTTTTCTGAGCGCTCTCTGGGAGGACTGTGCCT 1367
DB 239 GGATGGTGAAGGGAATTCAGTTTACTTTTCTGAGCGCTCTCTGGGAGGACTGTGCCT 298

QY 1368 TGCGTGGGTGGTGTGGGGTCCCAAGTTCCTGGGTGTTCTCAGTTGGAAGCAAGAGCC 1427
DB 299 TGCGTGGGTGGTGTGGGGTCCCAAGTTCCTGGGTGTTCTCAGTTGGAAGCAAGAGCC 358

QY 1428 AACTCAGGGGTGAGGGTCCCAAGTTCCTGGGTGTTCTCAGTTGGAAGCAAGAGCTGTGAG 1487
DB 359 AACTCAGGGGTGAGGGTCCCAAGTTCCTGGGTGTTCTCAGTTGGAAGCAAGAGCTGTGAG 418

QY 1488 GGCAGGGGTGGAGGGTGTGAGCTGAAGAAAGCAGAGTTCGCGGGCAGGTGGCT 1547
DB 419 GGCAGGGGTGGAGGGTGTGAGCTGAAGAAAGCAGAGTTCGCGGGCAGGTGGCT 478

QY 1548 CACGCTGTAAATCCAGCACTTTGGGAGGCGCAGGCGGCGAGATCACAGGTCAGGAGAT 1607
DB 479 CAYGCTGTAAATCCAGCACTTTAGGAGCGCGGCGGCGAGATCACAGGTCAGGAGAT 538

QY 1608 CGAGACCATCTCGGTGTAACCGGGGTGAACCCCGTCTCTACTAAATAATCAAAAAATCA 1667
DB 1608 CGAGACCATCTCGGTGTAACCGGGGTGAACCCCGTCTCTACTAAATAATCAAAAAATCA 1667

```

```
Db 419 GGCAGGGTGGAGGGTGTGAGACTGAAGAAAGCAGGAGTTCGCGGCACGGTGGCT 478
QY 1548 CACGCTCTAATCCAGACACTTTGGAGGCCAGAGCGGGCAGATCACAGGTTCAGGAGAT 1607
Db 479 CAYGCTCTAATCCAGACACTTTAGGAGGCCAGAGCGGGCAGATCACAGGTTCAGGAGAT 538
QY 1608 CGAGACCATCTGGCTAAACAGGGGTGAACCCCGTCTCTACTAATAAATACAAAAATCA 1667
Db 539 CGAGACCATCTGGCTAAACAGGGGTGAACCCCGTCTCTACTAATAAATACAAAAATCA 1667
QY 1668 GCCG 1671
Db 597 GCCG 600

RESULT 6
US-10-357-930-23318/c
; Sequence 23318, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23318
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1467, 1476, 1497
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23318

Query Match 11.9%; Score 337; DB 20; Length 1506;
Best Local Similarity 99.7%; Pred. No. 8.5e-84;
Matches 348; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 977 GGCACAGCTGTCTCTCAACGAGTTTGTGAAGTGCCCGTCGGGACAAAGTGGGTGATGAAG 1036
Db 1325 GGCACAGCTGTCTCTCAACGAGTTTGTGAAGTGCCCGTCGGGACAAAGTGGGTGATGAAG 1266
QY 1037 ATGCTGCAGATGGACATGAATCCACGACAGCTGGCTCGCTCAGCAGACGGAAGTGC 1096
Db 1265 ATGCTGCAGATGGACATGAATCCACGACAGCTGGCTCGCTCAGCAGACGGAAGTGC 1206
QY 1097 ATGTTCTGAGGAGTCTGGGGCCCTCCAGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1156
Db 1205 ATGTTCTGAGGAGTCTGGGGCCCTCCAGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1146
QY 1097 ATGTTCTGAGGAGTCTGGGGCCCTCCAGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1156
Db 1205 ATGTTCTGAGGAGTCTGGGGCCCTCCAGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1146
QY 1157 TAGGAGGGTCCCTGGCTCAGCTGCTCATGCCCACTCTTCCCTGGTGTGACTTCCTG 1216
Db 1205 ATGTTCTGAGGAGTCTGGGGCCCTCCAGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1146
QY 1157 TAGGAGGGTCCCTGGCTCAGCTGCTCATGCCCACTCTTCCCTGGTGTGACTTCCTG 1216
```

```
Db 1145 TAGAGGGT-CCCTGGCTCAGCTGCTCATGCCCACTCTTCCCTGGTGTGACTTCCTG 1087
QY 1217 GCACCCCTGTGCAGGGCTGAGTGGGGATGGGGAAGGCTGCTGGGTTTGAAGTGCCCA 1276
Db 1086 GCACCCCTGTGCAGGGCTGAGTGGGGATGGGGAAGGCTGCTGGGTTTGAAGTGCCCA 1027
QY 1277 CAGGTCATAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGGGAATTC 1325
Db 1026 CAGGTCATAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGGGAATTC 978

RESULT 7
US-10-357-930-29199/c
; Sequence 29199, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29199
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1467, 1476, 1497
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29199

Query Match 11.9%; Score 337; DB 20; Length 1506;
Best Local Similarity 99.7%; Pred. No. 8.5e-84;
Matches 348; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 977 GGCACAGCTGTCTCTCAACGAGTTTGTGAAGTGCCCGTCGGGACAAAGTGGGTGATGAAG 1036
Db 1325 GGCACAGCTGTCTCTCAACGAGTTTGTGAAGTGCCCGTCGGGACAAAGTGGGTGATGAAG 1266
QY 1037 ATGCTGCAGATGGACATGAATCCACGACAGCTGGCTCGCTCAGCAGACGGAAGTGC 1096
Db 1265 ATGCTGCAGATGGACATGAATCCACGACAGCTGGCTCGCTCAGCAGACGGAAGTGC 1206
QY 1097 ATGTTCTGAGGAGTCTGGGGCCCTCCAGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1156
Db 1205 ATGTTCTGAGGAGTCTGGGGCCCTCCAGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1146
QY 1157 TAGGAGGGTCCCTGGCTCAGCTGCTCATGCCCACTCTTCCCTGGTGTGACTTCCTG 1216
Db 1145 TAGGAGGGT-CCCTGGCTCAGCTGCTCATGCCCACTCTTCCCTGGTGTGACTTCCTG 1087
QY 1217 GCACCCCTGTGCAGGGCTGAGTGGGGATGGGGAAGGCTGCTGGGTTTGAAGTGCCCA 1276
```

```
Db 1086 GCACCCCTGTGAGGGCTGAGTGGGGATGGGAAGGCTGCTGGGTTTGAAGTGGCCAA 1027
QY 1277 CAGGGCATAGTCCATTTTGGAGAGTCCCTTGGGATCGTGAAGGGAATTC 1325
Db 1026 CAGGGCATAGTCCATTTTGGAGAGTCCCTTGGGATCGTGAAGGGAATTC 978

RESULT 8
US-10-266-103-39
; Sequence 39, Application US/10266103
; Publication No. US20040242510A1
; GENERAL INFORMATION:
; APPLICANT: Markowitz, Sanford D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING COLON CANCERS
; FILE REFERENCE: CWRU-P01-008
; CURRENT APPLICATION NUMBER: US/10/266,103
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/327537
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 62520
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-266-103-39

Query Match 9.8%; Score 277.2; DB 20; Length 62520;
Best Local Similarity 88.6%; Pred. No. 5e-66;
Matches 312; Conservative 0; Mismatches 38; Indels 2; Gaps 1;

QY 1518 AAAAGGCAGGAGTTCGGCGCAGGTGGCTCAGCGCTGTAATCCCAAGCACTTTGGGAGGC 1577
Db 29737 AAAATACAGATGGCGCGCGCGGTGGCTCAGCGCTGTAATCCCAAGCACTTTGGGAGGC 29796
QY 1578 CGAGCGGCGAGATCACAGAGTTCAGAGATCGAGACCATCTGGCTGTAACACGGGGTGAAA 1637
Db 29797 CGAGCGGCGGATCACAGAGTTCAGAGATTCAGAGACCATCTGGCTGTAACAC--GGTGAAA 29854
QY 1638 CCCCGTCTCTACTAAATAACAAAATACAGCGGTGAGGTGGCGGCGCTGTAGTCC 1697
Db 29855 CCCCGTCTCTACTAAATAACAAAATTAGCCGGGAGGTGGCGGCGCTGTAGTCC 29914
QY 1698 CAGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCCAAGGGCGCGAGCCTACAG 1757
Db 29915 CAGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCCAAGGGCGCGAGCCTACAG 29974
QY 1758 TGAGCGGAGATTGGCGCACTGCACCTCAGCGCTGAGAGAGTGGAGAGTCCGTTCTCAAAA 1817
Db 29975 TGAGCGGAGATTGGCGCACTGCACCTCAGCGCTGAGAGAGTGGAGAGTCCGTTCTCAAAA 30034
QY 1818 AAAAAAAGAAAGAAAGAGGAGGAGTTTGGGGGCGAGGGGAGCA 1869
Db 30035 AAAAAAAGAAAGAAAGAGGAGGAGTTTGGGGGCGAGGGGAGCA 30086

RESULT 9
US-09-764-855-328/c
; Sequence 328, Application US/09764855
; Patent No. US20020119919A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA110
; CURRENT APPLICATION NUMBER: US/09/764,855
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 328
; LENGTH: 32152
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-764-855-328

Query Match 9.7%; Score 275; DB 9; Length 32152;
Best Local Similarity 88.1%; Pred. No. 1.4e-65;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

QY 1486 GAGGCAGGGTGGGAGGGTGTGAGACTGAAGAAAGGAGGAGTTCGCGGCGACGGTGG 1545
Db 18284 GAGGCGGGTGGATCATGAAGTGAAGAGATCAAGACCATCTCTGTTGGCGGCGCGGTGG 18225
QY 1546 CTCAGCGCTGTAAATCCCAAGCACTTTGGGAGGCGGAGGCGGAGATCAGAGGTCAAGGAG 1605
Db 18224 CTCAGCGCTGTAAATCCCAAGCACTTTGGGAGGCGGAGGCGGCGGATCAGAGGTCAAGGAG 18165
QY 1606 ATCGAGACCATCTCGCTAAACACGGGGTGAACACCCCGTCTCTACTAAAAATACAAAAAT 1665
Db 18164 ATCGAGACCATCTCGCTAAACAC--GGTGAACCCCGTCTCTACTAAAAATACAAAAAT 18107
QY 1666 CAGCGGGTGAAGTGGCGGCGCTGTAGTCCAGTACTCAGGAGGCTGAGGCAAGAGA 1725
Db 18106 TAGCCGGGCGAGGTGGCGGCGCTGTAGTCCAGTACTCAGGAGGCTGAGGCAAGAGA 18047
QY 1726 ATGGCGTGAACCCCAAGGGCGAGCCTACAGTGAAGCGGAGTTCGCGCACTGCCA 1785
Db 18046 ATGGCGTGAACCCCAAGGGCGAGCCTGTGCACTGAGCGGAGTTCGCGCACTGCCA 17987
QY 1786 GCCTGGCAGACAGTGCAGACTCCGCTCTCAAAAAAAGAAAGAAAGAAA 1838
Db 17986 GCCTGGCAGACAGCAGACTCCGCTCTCAAAAAAAGAAAGAAAGAAA 17934

RESULT 10
US-09-764-872-518
; Sequence 518, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT APPLICATION NUMBER: US/09/764,872
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 518
; LENGTH: 32152
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-872-518

Query Match 9.7%; Score 275; DB 10; Length 32152;
Best Local Similarity 88.1%; Pred. No. 1.4e-65;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

QY 1486 GAGGCAGGGTGGGAGGGTGTGAGACTGAAGAAAGGAGGAGTTCGCGGCGACGGTGG 1545
Db 13869 GAGGCGGGTGGATCATGAAGTGAAGAGATCAAGACCATCTCTGTTGGCGGCGCGGTGG 13928
QY 1546 CTCAGCGCTGTAAATCCCAAGCACTTTGGGAGGCGGAGGCGGAGATCAGAGGTCAAGGAG 1605
Db 13929 CTCAGCGCTGTAAATCCCAAGCACTTTGGGAGGCGGAGGCGGCGGATCAGAGGTCAAGGAG 13988
QY 1606 ATCGAGACCATCTCGCTAAACACGGGGTGAACACCCCGTCTCTACTAAAAATACAAAAAT 1665
Db 13989 ATCGAGACCATCTCGCTAAACAC--GGTGAACCCCGTCTCTACTAAAAATACAAAAAT 14046
QY 1666 CAGCGGGTGAAGTGGCGGCGCTGTAGTCCAGTACTCAGGAGGCTGAGGCAAGAGA 1725
Db 14047 TAGCCGGGCGAGGTGGCGGCGCTGTAGTCCAGTACTCAGGAGGCTGAGGCAAGAGA 14106
QY 1726 ATGGCGTGAACCCCAAGGGCGGAGCCTACAGTGAAGCGGAGTTCGCGCACTGCCA 1785
```

Db 14107 ATGGCGTGAACCCAGGGCGGAGCTGCACTGAGCCGAGATTGCGCCACTGCCTCCA 14166
QY 1786 GCCTGACGACAGTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 1838
Db 14167 GCCTGGCGACGAGAGTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 14219

RESULT 11
US-10-072-349-328/c
; Sequence 328, Application US/10072349
; Publication No. US20030054420A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P110C1
; CURRENT APPLICATION NUMBER: US/10/072,349
; CURRENT FILING DATE: 2002-02-11
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 328
; LENGTH: 32152
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-349-328

Query Match 9.7%; Score 275; DB 14; Length 32152;
Best Local Similarity 88.1%; Pred. No. 1.4e-65;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;
QY 1486 GAGCGGGGTGGAGGGTGTGAGACTGAAGAAAGGAGGAGTTCGCGGCGACCGTGG 1545
Db 18284 GAGCGGGTGGATCATGAAGTGAAGAGATCAAGACCATCTGTTGGCGGCGCGTGG 18225
QY 1546 CTCAGCCCTGTAATCCAGCACTTTGGAGGCGGAGGCGGAGATCAGAGTCAGAGT 1605
Db 18224 CTCAGCCCTGTAATCCAGCACTTTGGAGGCGGAGGCGGAGATCAGAGTCAGAG 18165
QY 1606 ATCGAGACCATCTCGCTAACAACGGGGTGAACCCGCTCTACTAAAAATACAAAA 1665
Db 18164 ATCGAGACCATCTCGCTAACAAC--GGTGAACCCGCTCTACTAAAAATACAAAA 18107
QY 1666 CAGCGGGTGAAGTGGCGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGGCGAAGAG 1725
Db 18106 TAGCGGGCGAGTGGCGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGGCGAAGAG 18047
QY 1726 ATGGCGTGAACCCAGGGGCGAGCTTACAGTGAAGCGAGATTGCGCCACTGCCTCCA 1785
Db 18046 ATGGCGTGAACCCAGGGGCGAGCTTACAGTGAAGCGAGATTGCGCCACTGCCTCCA 17987
QY 1786 GCCTGACGACAGTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 1838
Db 17986 GCCTGGCGACGAGACTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 17934

RESULT 12
US-10-087-192-1240
; Sequence 1240, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; Prior Application removed - See file Wrapper or PALM
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1240
; LENGTH: 129042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1240
Query Match 9.7%; Score 275; DB 13; Length 129042;
Best Local Similarity 88.1%; Pred. No. 3.1e-65;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;
QY 1486 GAGCGGGGTGGAGGGTGTGAGACTGAAGAAAGGAGGAGTTCGCGGCGACCGTGG 1545
Db 123352 GAGCGGGTGGATCATGAAGTGAAGAGATCAAGACCATCTGTTGGCGGCGCGTGG 123411
QY 1546 CTCAGCCCTGTAATCCAGCACTTTGGAGGCGGAGGCGGAGATCAGAGTTCAGGAG 1605
Db 123412 CTCAGCCCTGTAATCCAGCACTTTGGAGGCGGAGGCGGAGATCAGAGTTCAGGAG 123471
QY 1606 ATCGAGACCATCTCGCTAACAACGGGGTGAACCCGCTCTACTAAAAATACAAAA 1665
Db 123472 ATCGAGACCATCTCGCTAACAAC--GGTGAACCCGCTCTACTAAAAATACAAAA 123529
QY 1666 CAGCGGGTGAAGTGGCGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGGCGAAGAG 1725
Db 123530 TAGCGGGCGAGTGGCGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGGCGAAGAG 123589
QY 1726 ATGGCGTGAACCCAGGGGCGGAGCTTACAGTGAAGCGAGATTGCGCCACTGCCTCCA 1785
Db 123590 ATGGCGTGAACCCAGGGGCGGAGCTTACAGTGAAGCGAGATTGCGCCACTGCCTCCA 123649
QY 1786 GCCTGACGACAGTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 1838
Db 123650 GCCTGGCGACAGGAGACTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 123702

RESULT 13
US-10-242-355-840/c
; Sequence 840, Application US/10242355
; Publication No. US20030235831A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC003C1
; CURRENT APPLICATION NUMBER: US/10/242,355
; CURRENT FILING DATE: 2002-09-13
; Prior Application removed - See file Wrapper or PALM
; NUMBER OF SEQ ID NOS: 1267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 840
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-355-840

Query Match 9.7%; Score 274.6; DB 17; Length 335;
Best Local Similarity 92.0%; Pred. No. 1.5e-66;
Matches 301; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

QY 1513 TGAAGAAAAAGGAGGAGTTGCCGGGACCGGTGGCTCACGCCCTGTAAATCCAGCAGCTTTGG 1572
DB 334 TTATATAATGCGAGGGAAGCGCGGGCGGTGGCTCACGCCCTGTAAATCCAGCAGCTTTGG 275

QY 1573 GAGGCGGAGCGGCGGAGATCAGAGGTGAGGAGATCGAGACCATCTGGCTTAACACGGGG 1632
DB 274 GAGGCGGAGCGGCGGTGGATCAGAGGTGAGGAGATCGAGACCATCTGGCTTAACAC--GG 217

QY 1633 TGAACCCCGCTCTCTACTTAAAAATACAAAAATCAGCCGGGTGAGGTGGCGGGCGGCTGT 1692
DB 216 TGAACCCCGCTCTCTACTTAAAAATACAAAAATTTAGCCGGGGGAGGTGGCGGGCGGCTGT 157

QY 1693 AGTCCCGAGCTACTCAGGAGGCTGAGGCAAGAGAAATGGCGGTGAACCCAGGGGGCGGAGCC 1752
DB 156 AGTCCCGAGCTACTCGGAGGCTGAGGCAAGAGAAATGGCGGTGAACCCAGGGGGCGGAGCC 97

QY 1753 TACAGTGAGCGAGATTTGCCCACTGCACTCCAGCCTGGAGCAGAGTGGAGACTCCGCTCTC 1812
DB 96 TCGAGTGAGCGGAGATTTGCCCACTGCACTCCAGCCTGGGCGAGAGACTCCGCTCTC 37

QY 1813 AAAAAAAGAAAAAAGAAAAA 1839
DB 36 AAAAAAAGAAAAAAGAAAAA 10

RESULT 14

US-10-087-192-544
; Sequence 544, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 544
; LENGTH: 238484
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-544

Query Match 9.7%; Score 273.4; DB 13; Length 238484;
Best Local Similarity 90.2%; Pred. No. 1.2e-64;
Matches 304; Conservative 0; Mismatches 31; Indels 2; Gaps 1;

QY 1503 GTGTTGAGACTGAAGAAAGGAGGAGTTGCCGGGACCGGTGCTCACGCCCTGTAAATCC 1562
DB 38352 GTGGCAAAAGTGTGTAAGAAATGATGGCGCGGGCGGTGGCTCACGCCCTGTAAATCC 38411

QY 1563 AGCATTTCGGAGGCGGCGGAGATCAGAGGTGAGGAGATCGAGACCATCTTGGC 1622
DB 38412 AGCATTTCGGAGGCGGCGGAGATCAGAGGTGAGGAGATCGAGACCATCTTGGC 38471

QY 1623 TAAACGGGTGAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGGTGAGGTGGC 1682
DB 38472 TAAAC--GGTGAACCCCTGTCTCTACTAAAAATACAAAAATTAGCCGGGCGAGGTGC 38529

QY 1683 GGGCGCTGTAGTCCAGCTACTCAGGAGGTGAGGCAAGAGAAATGGCGTGAACCCAGG 1742
DB 38530 GGGCGCTGTAGTCCAGCTACTCAGGAGGTGAGGCAAGAGAAATGGCGTGAACCCAGG 38589

QY 1743 GGGCGGAGCCTACAGTGAAGCGAGATTGGCCCACTGCAGCTCCAGCCTGACACAGTGA 1802
DB 38590 GGGCGGAGCCTGCTGAGTGAAGCGAGATTGGCCCACTGCAGCTCCAGCCTGACACAGTGA 38649

QY 1803 ACTCCGTCTCAAAAAAAGAAAAAAGAAAAA 1839
DB 38650 ACTCCGTCTCAAAAAAAGAAAAAAGAAAAAATATAAATA 38686

RESULT 15

US-10-220-510-5/c
; Sequence 5, Application US/10220510
; Publication No. US20030190637A1
; GENERAL INFORMATION:
; APPLICANT: Hovnanian, Alain
; APPLICANT: Chavanas, Stephanie
; APPLICANT: Cookson, William
; APPLICANT: Moffat, Miriam
; APPLICANT: Walley, Andrew
; TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR NETHERTON'S DISEASE
; FILE REFERENCE: I00317.70008.US
; CURRENT APPLICATION NUMBER: US/10/220,510
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: GB 0005098.9
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: GB 0005229.0
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7995
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7994)..(7995)
; OTHER INFORMATION: n = a, c, g or t/u
US-10-220-510-5

Query Match 9.7%; Score 272.6; DB 16; Length 7995;
Best Local Similarity 94.8%; Pred. No. 3.2e-65;
Matches 293; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

QY 1532 GCCGGGCGAGTGGCTCACGCTGTAAATCCAGCAGCTTTGGAGGCGCGGCGGAGAT 1591
DB 3681 GCCGGGCGGTTGGCTCACGCTGTAAATCCAGCAGCTTTGGAGGCGCGGCGGAGAT 3622

QY 1592 CACGAGGTGAGGAGATCGAGACCATCTGGCTTAACACGGGGTGAACCCCGTCTCTACTA 1651
DB 3621 CACGAGGTGAGGAGATCGAGACCATCTGGCTTAACAC--GGTGAACCCCGTCTCTACTA 3564

QY 1652 AAATACAAAAAATCAGCGGTGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAG 1711
DB 3563 AAATACAAAAAATTAGCGGGCGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAG 3504

QY 1712 GCTGAGGCAAGAGAAATGGCGTGAACCCAGGCGGGCGGAGCCTACAGTGAAGCGGAGTTGC 1771
DB 3503 GCTGAGGCAAGAGAAATGGCGTGAACCCAGGCGGGCGGAGCCTGCACTGAGCGAGATTGC 3444

QY 1772 GCCACTGCACTCCAGCTCGGAGCAGCAGTGAAGTCCGTCTCAAAAAAAGAAAAAAGA 1831
DB 3443 GCCACTGCACTCCAGCTCGGAGCAGCAGTGAAGTCCGTCTCAAAAAAAGAAAAAAGA 3384

QY 1832 AAAGAAAAAG 1840
DB 3383 AAAAAAAG 3375

Search completed: June 29, 2005, 12:01:59
Job time : 1107.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 15:06:07 ; Search time 6070.87 Seconds
(without alignments)
17687.639 Million cell updates/sec

Title: US-08-731-499-6
Perfect score: 2821
Sequence: 1 ATCTTAGCGCACACCTG.....AGGGTTCAGACGACCTG 2821

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:

1: gb_est1:*
2: gb_est2:*
3: gb_hgc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	766	27.2	779	5	EX111942
2	615.4	21.8	668	4	BM690083
3	560.8	19.9	572	4	BM690003
4	555	19.7	632	5	BQ637863
5	537	19.0	559	4	BM662840
6	524.6	18.6	546	5	BQ447506
7	520.8	18.5	672	2	AW967092
8	500.2	17.7	534	4	BM661803
9	486	17.2	486	2	BE220617
10	466.8	16.5	471	1	AI377509
11	464.8	16.4	489	4	BM688117
12	461.8	16.4	489	1	AI424326
13	456.4	16.2	475	4	BM704154
14	448.8	15.9	455	2	AW020233
15	436	15.5	437	4	BM686135
16	418.2	14.8	439	1	AI306331
17	408.6	14.5	442	4	BM672390
18	397.2	14.1	407	4	BM701070
19	380.8	13.5	399	4	BI293114
20	376.4	13.3	511	8	AQ767124
21	374	13.3	407	7	H40682
22	363	12.9	378	2	AW136419
23	335.8	11.9	404	1	AA308846
24	325	11.5	740	7	W27113

C	25	324.8	11.5	456	7	W26445	W26445 29f7 Human
	26	321.8	11.4	608	7	W27002	W27002 19h12 Human
	27	319	11.3	782	7	W28792	W28792 52f3 Human
C	28	314.4	11.1	579	7	W27379	W27379 28c1 Human
	29	311.8	11.1	331	1	AA317761	AA317761 EST19742
C	30	311	11.0	829	7	W26485	W26485 30c1 Human
	31	302.6	10.7	483	7	W27135	W27135 22h4 Human
C	32	296.4	10.5	843	7	W28227	W28227 43h1 Human
	33	293.4	10.4	342	1	AA317311	AA317311 EST19308
	34	276	9.8	760	7	W27109	W27109 20d3 Human
	35	267.6	9.5	532	8	AQ184326	AQ184326 HS 3220.A
	36	266	9.4	267	1	AA378172	AA378172 EST90835
	37	264.8	9.4	5670	3	HS8004253	AL832942 Homo sapi
C	38	264.4	9.4	311	5	BQ429229	BQ429229 AGENCOURT
	39	260.2	9.2	295	5	BUT727375	BU727375 UI-E-CQ1-
	40	259.6	9.2	498	8	AQ199645	AQ199645 RPI111-62
	41	258.6	9.2	570	7	CR542495	CR542495 DKF2p459J
	42	258.6	9.2	630	7	CR542494	CR542494 DKF2p459J
	43	256.4	9.1	409	8	AQ627863	AQ627863 CTBTI-E1-
C	44	256.2	9.1	334	5	BX493558	BX493558 DKF2p781M
C	45	252.2	8.9	297	4	BM663617	BM663617 UI-E-CK0-

ALIGNMENTS

RESULT 1
BX111942
LOCUS BX111942 779 bp mRNA linear EST 07-FEB-2003
DEFINITION BX111942 Soares retina N2b4HR Homo sapiens cDNA clone
IMAGP998H09360 ; IMAGE:190904, mRNA sequence.
ACCESSION BX111942
VERSION BX111942.1 GI:27878691
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 779)
AUTHORS Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
TITLE Human Unigeneset - RZPD3
JOURNAL Unpublished (2003)
COMMENT Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGP998H09360
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.

FEATURES

Location/Qualifiers
1..779
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGP998H09360 ; IMAGE:190904"
/sex="male"
/tissue_type="retina"
/dev_stage="55 year old"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares retina N2b4HR"
/note="Organ: eye; Vector: p7713D (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not 1 - oligo(dT) primer [5"

TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTTTTTTTTTTTTTTTTT 3'}.
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). The retinas were obtained from a 55 year old
Caucasian and total cellular poly(A)+ RNA was extracted 6
hrs after their removal. The retina RNA was kindly
provided by Roderick R. McInnes M.D. Ph.D. from the
University of Toronto. Library constructed by Bento
Soares and M.Fatima Bonaldo. "

ORIGIN		Query Match	27.2%;	Score 766;	DB 5;	Length 779;
		Best Local Similarity	99.6%;	Pred. No. 2.3e-126;		
		Matches 777;	Conservative 0;	Mismatches 2;	Indels 1;	Gaps 1;
QY	1577	CCGAGCGCGGCGAGATACGAGGTTCAGGAGATCGAGACCAATCTGGCTAACCGGGGTGAA	1636			
Db	1	CCGAGCGCGGCGAGAT-NCGAGGTTCAGGAGATCGAGACCAATCTGGCTAACCGGGGTGAA	59			
QY	1637	ACCCGCTCTCTACTAAATATACAAATATCAGCGGCTGAGTGGCGGCGCTGTATCT	1696			
Db	60	ACCCGCTCTCTACTAAATATACAAATATCAGCGGCTGAGTGGCGGCGCTGTATCT	119			
QY	1697	CCAGCTACTCAGAGGCTGAGGCAAGAGATGGCGTGAACCCAGCGGGCCGAGCTACA	1756			
Db	120	CCAGCTACTCAGAGGCTGAGGCAAGAGATGGCGTGAACCCAGCGGGCCGAGCTACA	179			
QY	1757	GTGAGCCGAGATTGGCCCATCTGACTCCAGCTGGAAGCAGTGAAGTCCGTCTCAAAA	1816			
Db	180	GTGAGCCGAGATTGGCCCATCTGACTCCAGCTGGAAGCAGTGAAGTCCGTCTCAAAA	239			
QY	1817	AAAAAAAAGAAAAGAAAAGGAGAGATTTTGGGGGGGAGGGGCGAGCAATATTC	1876			
Db	240	AAAAAAAAGAAAAGAAAAGGAGAGATTTTGGGGGGGAGGGGCGAGCAATATTC	299			
QY	1877	TATAACTTCGGGATGCTGAGGGGCTTCATGGGAGGAGCCCTGGCTCTCTCCCAA	1936			
Db	300	TATAACTTCGGGATGCTGAGGGGCTTCATGGGAGGAGCCCTGGCTCTCTCCCAA	359			
QY	1937	GGCATCTCTCAGTGGTGTCAACAGGAAAAATGGCAGCAATATGCTGAGGCTGTGT	1996			
Db	360	GGCATCTCTCAGTGGTGTCAACAGGAAAAATGGCAGCAATATGCTGAGGCTGTGT	419			
QY	1997	CTTTCTGCTTTGAAAGGCTGAGTGTACTTAAAGGAGCTTTTCACTCTGCTGGGTG	2056			
Db	420	CTTTCTGCTTTGAAAGGCTGAGTGTACTTAAAGGAGCTTTTCACTCTGCTGGGTG	479			
QY	2057	CTGCTCTGGGACCCCTCTGCTGCAACCCACCACTCCCCCAACAATCTCTCTTCATCC	2116			
Db	480	CTGCTCTGGGACCCCTCTGCTGCAACCCACCACTCCCCCAACAATCTCTCTTCATCC	539			
QY	2117	ATATCCCCAGTATGAGACCTTCCACAATCTCCAGCCATTAAGCTGAATGTTCTTTTAA	2176			
Db	540	ATATCCCCAGTATGAGACCTTCCACAATCTCCAGCCATTAAGCTGAATGTTCTTTTAA	599			
QY	2177	GGATGAGAAAATCTCTGCTCTGCAAGAAATGGGGAGCTGTTCAGTGGGATGT	2236			
Db	600	GGATGAGAAAATCTCTGCTCTGCAAGAAATGGGGAGCTGTTCAGTGGGATGT	659			
QY	2237	GGGCTGGGCTTGGCTTCTAACTGCTGTGACCCAGAGCAGCTTCTCTCCCTTAAC	2296			
Db	660	GGGCTGGGCTTGGCTTCTAACTGCTGTGACCCAGAGCAGCTTCTCTCCCTTAAC	719			
QY	2297	TTGGTTATGCTTTGGCAGCAGTGAAGAGGTGGACTAGGCGAACAAGTTTGGATTAT	2356			
Db	720	TTGGTTATGCTTTGGCAGCAGTGAAGAGGTGGACTAGGCGAACAAGTTTGGATTAT	779			

RESULT 2
BM690083 668 bp mRNA linear EST 28-FEB-2002
LOCUS UI-E-CK1-abr-f-10-0-UI.r1 UI-E-CK1 Homo sapiens cDNA clone
DEFINITION

FEATURES		source
ACCESSION	BM690083	
VERSION	BM690083.1	GI:19003341
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	1 (bases 1 to 668)	
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.	
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery	
JOURNAL	Genome Res. 6 (9), 791-806 (1996)	
MEDLINE	97044477	
PUBMED	8889548	
COMMENT	Contact: Soares, MB Coordinated Laboratory for Computational Genomics University of Iowa 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: bento-soares@uiowa.edu Tissue Procurement: Dr. Gregg Hageman cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).	
Seq primer:	M13 Reverse.	
Location/Qualifiers	1..668	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="UI-E-CK1-abr-f-10-0-UI"	
	/tissue type="Retina Foveal and Macular"	
	/dev stage="adult"	
	/lab host="DH10B (Life Technologies) (T1 phage resistant)"	
	/clone lib="UI-E-CK1"	
	/note="Organ: eye; Vector: p773-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-CK1 is a normalized cDNA library containing the following tissue(s): Retina Foveal and Macular. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTCC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."	

ORIGIN

Query Match		21.8%;	Score 615.4;	DB 4;	Length 668;
Best Local Similarity		97.0%;	Pred. No. 1.6e-99;		
Matches 650;		Conservative 0;	Mismatches 12;	Indels 8;	Gaps 2;
QY	1861	GGGGCAGCAATAATCTATAACTTCGGGATGCTGAGGGCGTTTCATGGGAGGACCCCTG	1920		
Db	7	GGGGCAGCAATAATCTATAACTTCGGGATGCTGAGGGCGTTTCATGGGAGGACCCCTG	66		
QY	1921	GCCTCTCTCTCCCCAAGGCATCTCACAGTGGTGTCAAAGGAAAAATGGCAGCAATA	1980		
Db	67	GCCTCTCTCTCCCCAAGGCATCTCACAGTGGTGTCAAAGGAAAAATGGCAGCAATA	126		
QY	1981	CGCTCAGGCTGTGCTCTTCTGCGCTTTGAAAGGGTCAGCTGTACTTAAAGGGAGCTGTTT	2040		
Db	127	CGCTCAGGCTGTGCTCTTCTGCGCTTTGAAAGGGTCAGCTGTACTTAAAGGGAGCTGTTT	186		

```

QY 2041 CAGCTCTGCTGGTGTCTGCTGGGACCCCTGCTGCCAACCCACCACTCCCCAACAA 2100
Db 187 CAGCTCTGCTGGTGTCTGCTGGGACCCCTGCTGCCAACCCACCACTCCCCAACAA 246
QY 2101 TCCTCTCTTTCCATCCATATCCCCAGTATGGACCTTCCACAACTCCAGCCATTAAGCTG 2160
Db 247 TCCTCTCTTTCCATCCATATCCCCAGTATGGACCTTCCACAACTCCAGCCATTAAGCTG 306
QY 2161 AATGTTTCTTTAAAGGATGGAGAAACTTTCTGTCTCTGCTCTGGCAAGATTGGGGGAC 2220
Db 307 AATGTTTCTTTAAAGGATGGAGAAACT---TCTGTCTCTGGCAAGATTGGGGGAC 362
QY 2221 TCTTGACTGGGATGTGGGCTGGCTGCTTCACTGCTGTGACCCAGACAGCCCA 2280
Db 363 TGTGTGACTGGGATGTGGGCTGGCTGCTTCACTGCTGTGACCCAGACAGCCCA 422
QY 2281 CTTCTCTCCCTAACTCTGTTATGCTTGGGACACAGTGACGCTGGACTAGGCCA 2340
Db 423 CTTCTCTCCCTAACTCTGTTATGCTTGGGACACAGTGACGCTGGACTAGGCCA 482
QY 2341 ACAGTTTTCGATATGTTTGTATGATGTGGAATTAATTTTGTATATAAATCTTAT 2400
Db 483 ACAGTTTTCGATATGTTTGTATGATGTGGAATTAATTTTGTATATAAATCTTAT 542
QY 2401 GTGTACCCCAATATAGAACTAGATTAAAGGAGCTCTCTGTTGAAAGGGAGCTG 2460
Db 543 GTGTACCCCAATATAGAACTAGATTAAAGGAGCTCTCTGTTGAAAGGGAGCTG 602
QY 2461 AGTACCTCTGGAACCTGGAGGACCTCTGAAAAAGCAAACTGAAAAACAGTGCCTGGG 2520
Db 603 AGTACCTCTGGAACCTGGAGGACCTCTGAAAAAGCAAACTGAAAAACAGTGCCTGGG 658
QY 2521 TCACTGTTTAC 2530
Db 659 TCACTGTTTAC 668

```

RESULT 3

```

BM690003
LOCUS
DEFINITION
  UI-E-CK1-abq-g-05-0-UI-ri UI-E-CK1 Homo sapiens cDNA clone
  UI-E-CK1-abq-g-05-0-UI 5', mRNA sequence.
BM690003
VERSION
KEYWORDS
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 572)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
JOURNAL
MEDLINE
PUBMED
COMMENT
  Genome Res. 6 (9), 791-806 (1996)
  9704447
  8889548
  Contact: Soares, MB
  Coordinated Laboratory for Computational Genomics
  University of Iowa
  375 Newton Road, 4156 MEBRP, Iowa City, IA 52242, USA
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: bento-soares@uiowa.edu
  Tissue Procurement: Dr. Gregg Hageman
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Researchers may obtain clones from Research
  Genetics (www.resgen.com).
  Seq primer: M13 Reverse.

```

FEATURES

```

source
  Location/Qualifiers
  1..572
  /organism="Homo sapiens"

```

RESULT 4

```

BQ637863
LOCUS
  BQ637863
  632 bp
  mRNA
  linear
  EST 15-JUL-2002

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CK1-abq-g-05-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-CK1 is a normalized cDNA library containing the
following tissue(s): Retina Foveal and Macular. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dr)18 tail. The sequence tag for this
library is GTCC. This library was created for the program,
Gene Discovery in the Visual System, supported by National
Eye Institute (NEI)."

```

ORIGIN

```

Query Match      19.9%; Score 560.8; DB 4; Length 572;
Best Local Similarity 99.3%; Pred. No. 9.1e-90;
Matches 562; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1845 GAGTTTTTGGGGGCGAGGGGCGAGCAGTAAATCTTAACTTCCGGGATGCTGAGGGCGTT 1904
Db 7 GGGTTTTTGGGGGCGAGGGGCGAGCAGTAAATCTTAACTTCCGGGATGCTGAGGGCGTT 66
QY 1905 CATGGGGAGGACCTTGGCCCTCTCTCTCCCAAGGCATCTCACACAGTGTGTCAACAGGA 1964
Db 67 CATGGGGAGGACCTTGGCCCTCTCTCTCCCAAGGCATCTCACACAGTGTGTCAACAGGA 126
QY 1965 AAAATGGCAGCAAAATACGCTGCAGGCTGTGGTCTTTCTGCCCTTTGAAAGGTCAGCTGTA 2024
Db 127 AAAATGGCAGCAAAATACGCTGCAGGCTGTGGTCTTTCTGCCCTTTGAAAGGTCAGCTGTA 186
QY 2025 CTTAAAGGAGCTGTTTCAGCTCTGCTGGGTGCTGCTCTGGGACCCCTGCTGCCAACCC 2084
Db 187 CTTAAAGGAGCTGTTTCAGCTCTGCTGGGTGCTGCTCTGGGACCCCTGCTGCCAACCC 246
QY 2085 ACCACTCCCCCAACAATCTCTCTTCCATCATATCCCCCAGTAGTGACCTTCCACAAC 2144
Db 247 ACCACTCCCCCAACAATCTCTCTTCCATCATATCCCCCAGTAGTGACCTTCCACAAC 306
QY 2145 TCCAGCCATAAGCTGAATGTTTCTCTTTAAAGGATGGAGAAACTTCTGTCTCTCTG 2204
Db 307 TCCAGCCATAAGCTGAATGTTTCTCTTTAAAGGATGGAGAAACTTCTGTCTCTCTG 366
QY 2205 GCAAGNATTTGGGGGACTGTTGACTGGGATGTGGCTGGGCTGGCTTCTTAAGTCTGTG 2264
Db 367 GCAAGNATTTGGGGGACTGTTGACTGGGATGTGGCTGGGCTGGCTTCTTAAGTCTGTG 426
QY 2265 TGACCCAGACAGCCACTTCTCTCCCTAACTTGGTTATGTCTTGGCAGCAGTAGTGAGC 2324
Db 427 TGACCCAGACAGCCACTTCTCTCCCTAACTTGGTTATGTCTTGGCAGCAGTAGTGAGC 486
QY 2325 AGGTCCGACTAGCGCAACAGTTTGGATTTATTTGTATTTAGATGGAATTAATTTTTTG 2384
Db 487 AGGTCCGACTAGCGCAACAGTTTGGATTTATTTGTATTTAGATGGAATTAATTTTTTG 546
QY 2385 TTATATAAATCTTATGTTAATCCCC 2410
Db 547 TTATATAAATCTTATGTTAATCCCC 572

```

```

DEFINITION hd15e08.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
ACCESSION BQ637863
VERSION BQ637863.1 GI:21762322
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 632)
AUTHORS Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A.,
Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE Expressed sequence tag analysis of human retina for the NEIBank
Project: Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
JOURNAL Mol. Vis. 8 (4), 196-204 (2002)
MEDLINE 22103461
PUBMED 12107411
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 15 row: e column: 08
Seq primer: M13RPI reverse primer (ABI).
Location/Qualifiers
1..632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hd15e08"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH108"
/clone_lib="Human Retina cDNA (Un-normalized,
unamplified): hd/he"
/note="organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the SuperScript Plasmid System full
details of which are contained in the manufacturer's
instruction manual (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I
primer-adaptor
[5'-PGACTAGTCTTAGATCGAGCGCGCC(T)15-3']. EST analysis
was performed on the unamplified library at the NIH
Intramural Sequencing Center (NISC)."
```

ORIGIN

```

Query Match 19.7%; Score 555; DB 5; Length 632;
Best Local Similarity 99.8%; Pred. No. 9.7e-89;
Matches 566; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 977 GGCACGCTCTCTGTGAACAGTTTGTGAAGTGGCCCGTCGGGACAGTGGGTGATGAAG 1036
DB 67 GGCACGCTCTCTGTGAACAGTTTGTGAAGTGGCCCGTCGGGACAGTGGGTGATGAAG 126

QY 1037 ATGCTGCAGATGGACATGATCCAGCAGCTGGCTGGCTCAGCAGAGCGGAAAGTGCC 1096
DB 127 ATGCTGCAGATGGACATGATCCAGCAGCTGGCTGGCTCAGCAGAGCGGAAAGTGCC 186

QY 1097 ATGTTCTGAGGAGTCTGGGGCCCTCCAGCACTCCAGGCTCCAGGTTTCCAGGGTAG 1156
DB 187 ATGTTCTGAGGAGTCTGGGGCCCTCCAGCACTCCAGGCTCCAGGTTTCCAGGGTAG 246

QY 1157 TAGGAGGGTCCCTGGCTCAGCGTGTCTATGACCACCTCTCCCTGGTGTGACTTCCTG 1216
DB 247 TAGGAGGGT-CCCTGGCTCAGCGTGTCTATGACCACCTCTCCCTGGTGTGACTTCCTG 305
```

```

QY 1217 GCACCCCTGTGCAGGCTGAGTGGGATGGGAAAGGCTGCTGGGTTTGAAGTGCCAA 1276
DB 306 GCACCCCTGTGCAGGCTGAGTGGGATGGGAAAGGCTGCTGGGTTTGAAGTGCCAA 365
QY 1277 CAGGCATAGTCCATTTTGGAGGAGTCCCTGGATGGTGAAGGAATTCAGTTACTTTTC 1336
DB 366 CAGGCATAGTCCATTTTGGAGGAGTCCCTGGATGGTGAAGGAATTCAGTTACTTTTC 425
QY 1337 CTGTTACGCGCTCTGGGAGGACTGCTGGCTGGGCTGGTGGTGGTGGGCTCCACAGT 1396
DB 426 CTGTTACGCGCTCTGGGAGGACTGCTGGCTGGGCTGGTGGTGGTGGGCTCCACAGT 485
QY 1397 TTCTGGGTGTTCTCAGTTGGAGCAAGCAACTGAGGGGTGAGGGTCCACAGACCAA 1456
DB 486 TTCTGGGTGTTCTCAGTTGGAGCAAGCAACTGAGGGGTGAGGGTCCACAGACCAA 545
QY 1457 ATCAGAAATGAGAACACAAAGACTGTTAGGAGCGAGGGTGGAGGGTGTTCAGACTGAA 1516
DB 546 ATCAGAAATGAGAACACAAAGACTGTTAGGAGCGAGGGTGGAGGGTGTTCAGACTGAA 605
QY 1517 GAAAGGAGGAGTTCGCGGCGCACCGT 1543
DB 606 GAAAGGAGGAGTTCGCGGCGCACCGT 632
```

RESULT 5

```

BM662840/c
LOCUS BM662840 559 bp mRNA linear EST 27-FEB-2002
DEFINITION UI-E-CK0-aam-b-12-0-UI.s1 UI-E-CK0 Homo sapiens cDNA clone
UI-E-CK0-aam-b-12-0-UI 3', mRNA sequence.
```

```

ACCESSION BM662840
VERSION BM662840.1 GI:18967679
```

```

KEYWORDS EST.
```

```

SOURCE Homo sapiens (human)
```

```

ORGANISM Homo sapiens
```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```

1 (bases 1 to 559)
```

```

Bonaldo,M.P., Lennon,G. and Soares,M.B.
```

```

Normalization and subtraction: two approaches to facilitate gene
```

```

discovery
```

```

Genome Res. 6 (9), 791-806 (1996)
```

```

97044477
```

```

8889548
```

```

Contact: Soares, MB
```

```

Coordinated Laboratory for Computational Genomics
```

```

University of Iowa
```

```

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
```

```

Tel: 319 335 8250
```

```

Fax: 319 335 9565
```

```

Email: bento-soares@uiowa.edu
```

```

Tissue Procurement: Dr. Gregg Hageman
```

```

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
```

```

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
```

```

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
```

```

Clone Distribution: Researchers may obtain clones from Research
```

```

Genetics (www.resgen.com).
```

```

Seq primer: M13 Forward
```

```

POLYA=Yes.
```

FEATURES

```

source
```

```

1..559
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CK0-aam-b-12-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK0"
/note="organ: eye; Vector: p7713-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-CK0 is a cDNA library containing the following
```

tissue(s): Retina Foveal and Macular. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTCC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG_TISSUE=Foveal and Macular Retina
TAG_LIB=UI-B-CX0
TAG_SEQ=GTCC"

ORIGIN

Query Match 19.0%; Score 537; DB 4; Length 559;
Best Local Similarity 99.1%; Pred. No. 1.6e-85;
Matches 540; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2095 CAACATCTCTCTTCCATCCATATCCCCAGTATGGACCTTCCACAACTCCAGCCAT 2154
DB 559 CAACATCTCTCTTCCATCCATATCCCCAGTATGGACCTTCCACAACTCCAGCCAT 500

QY 2155 AAGCTGAATGTTTCTCTTTAAAGGATGGAGAAAACCTTCTGTCTCTGGCAAGAATTG 2214
DB 499 AAGCTGAATGTTTCTCTTTAAAGGATGGAGAAAACCTTCTGTCTCTGGCAAGAATTG 440

QY 2215 GGGGACTGTGATGGGATTTGGGCTGGGCTTGGCTTCTTACTGCTGTGACCCAGA 2274
DB 439 GGGGACTGTGATGGGATTTGGGCTGGGCTTGGCTTCTTAACTGCTGTGACCCAGA 380

QY 2275 CAGCCACTTCTCTCCCTAACTTGGTATGTTGGCAGCAGTGACAGTGGACT 2334
DB 379 CAGCCACTTCTCTCCCTAACTTGGTATGTTGGCAGCAGTGACAGTGGACT 320

QY 2335 AGCGCAACAGTTTGGATPATTTGTTTTAGATGTGGAATTAATTTTTTGTATATAAC 2394
DB 319 AGCGCAACAGTTTGGATPATTTGTTTTAGATGTGGAATTAATTTTTTGTATATAAC 260

QY 2395 TCTTATGTATACCCCAATATAGAACTAGATTAAGAGGAGTCTCTCTGGTTGAAAGG 2454
DB 259 TCTTATGTATACCCCAATATAGAACTAGATTAAGAGGAGTCTCTCTGGTTGAAAGG 200

QY 2455 GAGCTGAGTACCTCTGGAACCTGGAGGACCTCTGAAAAAGCAACTGAAAAACAGTGC 2514
DB 199 GAGCTGAGTACCTCTGGAACCTGGAGGACCTCTGAAAAAGCAACTGAAAAACAGTGC 140

QY 2515 CTTGGGTCACTGTTACTCTTAAGACAGTTTTAAAGTGAGACCTGGAAAAACATTTGCTT 2574
DB 139 CTTGGGTCACTGTTACTCTTAAGACAGTTTTAAAGTGAGACCTGGAAAAACATTTGCTT 80

QY 2575 TACCTTGAATAGATAGTGTATGTTGGTATATAGAAATTAACCTATTAAACC 2634
DB 79 TACCTTGAATAGATAGTGTATGTTGGTATATAGAAATTAACCTATTAAACC 20

QY 2635 TGAGA 2639
DB 19 TGAAA 15

RESULT 6

LOCUS B0447506/c
DEFINITION UI-H-EUI-bae-m-11-0-UI.s1 NCI CGAP Ctl Homo sapiens cDNA clone
ACCESSION B0447506
VERSION B0447506.1 GI:21250611
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

B0447506 546 bp mRNA linear EST 29-MAY-2002
UI-H-EUI-bae-m-11-0-UI.s1 NCI CGAP Ctl Homo sapiens cDNA clone
UI-H-EUI-bae-m-11-0-UI 3', mRNA sequence.

REFERENCE
AUTHORS
TITLES
JOURNAL
COMMENT

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 546)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@iowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1..546
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EUI-bae-m-11-0-UI"
/tissue_type="Osteoarthritic Cartilage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP Ctl"
/note="Organ: Knee; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ctl is a normalized cDNA library containing the following tissue(s): Osteoarthritic Cartilage The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCACGCT.
TAG TISSUE=osteoarthritic cartilage
TAG_LIB=UI-H-EUI
TAG_SEQ=TGATCACGCT"

ORIGIN

Query Match 18.6%; Score 524.6; DB 5; Length 546;
Best Local Similarity 99.2%; Pred. No. 2.6e-83;
Matches 527; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2109 TTCCATCCATATCCCCAGTATGGACCTTCCACAACTCCAGCCATAAGCTGAATGTTTC 2168
DB 546 TTCCATCCATATCCCCAGTATGGACCTTCCACAACTCCAGCCATAAGCTGAATGTTTC 487

QY 2169 TCTTTAAGGATGGAGAAAACCTTCTGTCTCTGGCAAGAAATTGGGGGACTGTTGACT 2228
DB 486 TCTTTAAGGATGGAGAAAACCTTCTGTCTCTGGCAAGAAATTGGGGGACTGTTGACT 427

QY 2229 GGGATTGTGGGCTTGGCTTCTTAAGTGTGTGACCCCAAGACACCCACTTCTCCT 2288
DB 426 GGGATTGTGGGCTTGGCTTCTTAAGTGTGTGACCCCAAGACACCCACTTCTCCT 367

QY 2289 CCCTAACCTTGGTTATGTCTTGGCAGCAGTGAGCAGTCCGACTAGCGCAACAGTTTTT 2348
DB 366 CCCTAACCTTGGTTATGTCTTGGCAGCAGTGAGCAGTCCGACTAGCGCAACAGTTTTT 307

QY 2349 GGATTATTGTGTTTTAGATGTGGAATTAATTTTTTGTATATAAACTCTTATGTGTAACC 2408
DB 306 GGATTATTGTGTTTTAGATGTGGAATTAATTTTTTGTATATAAACTCTTATGTGTAACC 247

QY 2409 CCATATAGAACTAGATTAAAGGAGTCTCTCTGGTTGAAGGGGAGCTAGTACCTT 2468
DB 246 CCATATAGAACTAGATTAAAGGAGTCTCTCTGGTTGAAGGGGAGCTAGTACCTT 187

```

QY 2469 CTGGAACCTGGAGCACCTCTCTGAAAAAGCACAACCTGAAACAGTGCCCTGGGTCACTGTT 2528
    |||||||
Db 186 CTGGAACCTGGAGCACCTCTCTGAAAAAGCACAACCTGAAATCAAGTCCCTGGGTCACTGTT 127
QY 2529 ACTCCTATAAGACAGTCTTAAAGCTGAGACCTGGAAAAACAATTTGCTTTTACCTTTGAATAGAT 2588
    |||||||
Db 126 ACTCCTATAAGACAGTCTTAAAGCTGAGACCTGGAAAAACAATTTGCTTTTACCTTTGAATAGAT 67
QY 2589 AGGTTTTATGTTGGTATATAGAAATATAAACTTAACCTATTAAACCTGAGA 2639
    |||||||
Db 66 AGGTTTTATGTTGGTATATAGAAATATAAACTTAACCTATTAAACCTGAAA 16

RESULT 7
AW967092 672 bp mRNA linear EST 01-JUN-2000
LOCUS EST379166 MAGe resequences, MAGJ Homo sapiens cDNA, mRNA sequence.
ACCESSION AW967092
VERSION AW967092.1 GI:8156928
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 672)
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 236
Seq primer: Reverse.
FEATURES
    source
    1..672
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone_lib="MAGE resequences, MAGJ"
    /note="Vector: pBluescriptSKm"

ORIGIN
Query Match 18.5%; Score 520.8; DB 2; Length 672;
Best Local Similarity 99.4%; Pred. No. 1.2e-82; Indels 1; Gaps 1;
Matches 533; Conservative 0; Mismatches 2;

QY 2091 CCCCCAACAACTCCTCTCTTTCCATCCATATCCCCCAGTATGGACCTTCACAACTCCCGAG 2150
    |||||||
Db 1 CCCCCAACAACTCCTCTCTTTCCATCCATATCCCCCAGTATGGACCTTCACAACTCCCGAG 60
QY 2151 CCATAAGCTGAATGTTTCTCTTTAAAGGATGAGAAAACTTCTGTCTGTCTCTGCGCAAGA 2210
    |||||||
Db 61 CCATAAGCTGAATGTTTCTCTTTAAAGGATGAGAAAACTTCTGTCTGTCTGCGCAAGA 120
QY 2211 ATTGGGGGACTGTGTACTCGGGATTGGGCTGGGCTTCTTAAGTCTGTGTGACCC 2270
    |||||||
Db 121 ATTGGGGGACTGTGTACTCGGGATTGGGCTGGGCTTCTTAAGTCTGTGTGACCC 180
QY 2271 AAGACAGCCACTTCTCCTCCCTAAACCTTGGTTATGTCTTTGGCAGCAGTGGAGCGGTGG 2330
    |||||||
Db 181 AAGACAGCCACTTCTCCTCCCTAAACCTTGGTTATGTCTTTGGCAGCAGTGGAGCGGTGG 240
QY 2331 GACTAGGCGAACAGTTTGGATATATCTGTTTTAGATGTGAATATTTTTTGTATAT 2390
    |||||||
Db 241 GACTAGGCGAACAGTTTGGATATATCTGTTTTAGATGTGAATATTTTTTGTATAT 300

```

```

QY 2391 AAACCTCTTATGTGAACCCCAATATAGAAACCTAGATTAAAGGGAGTCTCTCTGGTTGAA 2450
    |||||||
Db 301 AAACCTCTTATGTGAACCCCAATATAGAAACCTAGATTAAAGGGAGTCTCTCTGGTTGAA 360
QY 2451 AGGGAGCTGAGTACCCCTCTGGAACTGGAGGACCTCTGAAAAAAGCAAACTGAAAAACA 2510
    |||||||
Db 361 AGGGAGCTGAGTACCCCTCTGGAACTGGAGGACCTCTGAAAAAAGCAAACTGAAAAACA 420
QY 2511 GTGCCCTGGGTCACTGTTTACTCTTATAAGACAGTTTAAAGTGAGACCTGGAAAAACATTT 2570
    |||||||
Db 421 GTGCCCTGGGTCACTGTTTACTCTTATAAGACAGTTTAAAGTGAGACCTGGAAAAACATTT 480
QY 2571 GCTTTACCTTGAATAGATAGTGTGTTTATCTTGGTATAT-AAGAAATAAACTAAACC 2625
    |||||||
Db 481 GCTTTACCTTGAATAGATAGTGTGTTTATCTTGGGATATAAGAAATAAACTAAACC 536

RESULT 8
BM661803 534 bp mRNA linear EST 27-FEB-2002
LOCUS UI-E-CK1-abr-f-10-0-UI.s1 UI-E-CK1 Homo sapiens cDNA clone
DEFINITION UI-E-CK1-abr-f-10-0-UI 3', mRNA sequence.
ACCESSION BM661803
VERSION BM661803.1 GI:18965842
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 534)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLYA=Yes.
FEATURES
    source
    1..534
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="UI-E-CK1-abr-f-10-0-UI"
    /tissue_type="Retina Foveal and Macular"
    /dev_stage="adult"
    /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
    /clone_lib="UI-E-CK1"
    /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
    modified polylinker; Site 1: EcoR I; Site 2: Not I;
    UI-E-CK1 is a normalized cDNA library containing the
    following tissue(s): Retina Foveal and Macular. The
    library was constructed according to Bonaldo, Lennon and
    Soares, Genome Research, 6:791-806, 1996. First strand
    cDNA synthesis was primed with an oligo-dT primer
    containing a Not I site. Double stranded cDNA was ligated
    to an EcoR I adaptor, digested with Not I, and cloned
    directionally into pT7T3-pac vector. The oligonucleotide
    used to prime the synthesis of first-strand cDNA contains
    a library tag sequence that is located between the Not I

```

site and the (dT)18 tail. The sequence tag for this library is GTCC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG TISSUE=Poveal and Macular Retina
TAG_LIB=UI-E-CK1
TAG_SEQ=GTCC"

ORIGIN	
Query Match	17.7%; Score 500.2; DB 4; Length 534;
Best Local Similarity	98.7%; Pred. No. 5.9e-79;
Matches	516; Conservative 0; Mismatches 3; Indels 4; Gaps 1;
QY	2117 ATATCCCGGAGTGGACCTTCACAACTCCAGCCATAGCTGAATGTTCTCTTTAA 2176
DB	534 ATATCCCGGAGTGGACCTTCACAACTCCAGCCATAGCTGAATGTTCTCTTTAA 475
QY	2177 GGATGAGAGAACTTCTGCTGCTCTGGCAAGATTGGGGACTCTGACCTGGGATTCT 2236
DB	474 GGATGAGAGAACTTCTGCTGCTCTGGCAAGATTGGGGACTCTGACCTGGGATTCT 419
QY	2237 GGCTGGGCTTGGCTTCTAACTCTGCTGACCAAGACAGCCACTTCTCTCCCTAAC 2296
DB	418 GGCTGGGCTTGGCTTCTAACTCTGCTGACCAAGACAGCCACTTCTCTCCCTAAC 359
QY	2297 TTGGTTATGCTTTGGCAGCAGTGGAGGTCGGAAGTGGGAGTGGGATTT 2356
DB	358 TTGGTTATGCTTTGGCAGCAGTGGAGGTCGGAAGTGGGAGTGGGATTT 299
QY	2357 GTGTTTTAGATGCGAATTTATTTTGTATATAACTCTTATGTAACCCCAATATA 2416
DB	298 GTGTTTTAGATGCGAATTTATTTTGTATATAACTCTTATGTAACCCCAATATA 239
QY	2417 GAACTAGATTTAAAGGGAGTCTCTCTGGTTGAAAGGGAGCTGAGTACCTCTGGA 2476
DB	238 GAACTAGATTTAAAGGGAGTCTCTCTGGTTGAAAGGGAGCTGAGTACCTCTGGA 179
QY	2477 GGAGGCACCTTGAAAAAGCAAACTGAAACAGTGCCTGGTCACTGTTACTCCTAT 2536
DB	178 GGAGGCACCTTGAAAAAGCAAACTGAAACAGTGCCTGGTCACTGTTACTCCTAT 119
QY	2537 AAGACAGTTTAAAGTGAAGCTGGAAACATTTGCTTTTACCTTGAATAGATGTTT 2596
DB	118 AAGACAGTTTAAAGTGAAGCTGGAAACATTTGCTTTTACCTTGAATAGATGTTT 59
QY	2597 ATGTTGGTATATAAGAAATAAACTAACTATTAACCTGAGA 2639
DB	58 ATGTTGGTATATAAGAAATAAACTAACTATTAACCTGAGA 16

RESULT 9
BE220617/c
LOCUS
DEFINITION
ht99c12.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3165238 3',
mRNA sequence.
ACCESSION
VERSION
BE220617.1 GI:8908020
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 486)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 450.

FEATURES
source
1. .486
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3165238"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu24"
/notes="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; plasmid DNA from the normalized
library NCI_CGAP_Lu5 was prepared, and as circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (clones
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match	17.2%; Score 486; DB 2; Length 486;
Best Local Similarity	100.0%; Pred. No. 2e-76;
Matches	486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	2143 ACTCCAGCATAAGCTGAATGTTCTCTTTAAAGGATGGAGAAACTTCTGCTGCTC 2202
DB	486 ACTCCAGCATAAGCTGAATGTTCTCTTTAAAGGATGGAGAAACTTCTGCTGCTC 427
QY	2203 TGGCAAGATTGGGGACTGTTGACCTGGGATTTGCTGGGCTGGCTTCTTAACCTG 2262
DB	426 TGGCAAGATTGGGGACTGTTGACCTGGGATTTGCTGGGCTGGCTTCTTAACCTG 367
QY	2263 TGTGACCCAGAGCAGCCACTTCTCTCCCTAACCTTTGGTTATGCTTTGGCAGCAGTGA 2322
DB	366 TGTGACCCAGAGCAGCCACTTCTCTCCCTAACCTTTGGTTATGCTTTGGCAGCAGTGA 307
QY	2323 GCAGGTCGAGTAGGGGAAACAGTTTGGATTTATGTTTATAGATGTGAATTTT 2382
DB	306 GCAGGTCGAGTAGGGGAAACAGTTTGGATTTATGTTTATAGATGTGAATTTT 247
QY	2383 TGTATATAAACTTTATGTTAAACCCCAATATAGAACTAGATTAAAGGGAGTCTCTC 2442
DB	246 TGTATATAAACTTTATGTTAAACCCCAATATAGAACTAGATTAAAGGGAGTCTCTC 187
QY	2443 TGGTTGAAAGGGAGCTGAGTACCTCTCGAACTGGAGGCACCTCTGAAAAAGCAAACT 2502
DB	186 TGGTTGAAAGGGAGCTGAGTACCTCTCGAACTGGAGGCACCTCTGAAAAAGCAAACT 127
QY	2503 GAAACACAGTGCCTGGGTCACCTGTTTACCTCTATATAAGACAGTTTAAAGTGAGACCTGGAA 2562
DB	126 GAAACACAGTGCCTGGGTCACCTGTTTACCTCTATATAAGACAGTTTAAAGTGAGACCTGGAA 67
QY	2563 AAACATTTCTTTACCTTGAATAGATGCTTTTATGTTGGTATATAAGAAATAAACTA 2622
DB	66 AAACATTTCTTTACCTTGAATAGATGCTTTTATGTTGGTATATAAGAAATAAACTA 7
QY	2623 ACCTAT 2628
DB	6 ACCTAT 1

RESULT 10
AI377509/c
LOCUS
DEFINITION
AI377509 471 bp mRNA linear EST 18-MAR-1999
tc37d08.x1 Soares total fetus_Nb2HF8_9w Homo sapiens cDNA clone
IMAGE:2066799 3' similar to contains element MER22 repetitive
element i, mRNA sequence.

ACCESSION AI377509
 VERSION AI377509.1 GI:4187362
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 471)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps@mail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 816 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 428.
 Location/Qualifiers
 FEATURES
 source
 1..471
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2066799"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 /clone_lib="Soares total_fetus Nb2HF8_9w"
 /note="Vector: pT730-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGGAGCGCCCTTAATTTTCTTTTCTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaudo. "
 ORIGIN
 Query Match 16.5%; Score 466.8; DB 1; Length 471;
 Best Local Similarity 99.4%; Pred. No. 5.4e-73;
 Matches 468; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2169 TCTTTAAAGGATGGAGAAAACCTTCTGCTCTCTGGCAAGAAATGGGGAGCTGTGACT 2228
 DB 471 TCTTTAAGGATGGAGAAACCTTCTGCTCTCTGGCAAGAAATGGGGAGCTGTGACT 412
 QY 2229 GGGATTGTGGGCTGGGCTTGGCTTCTAACTGCTGTGTGACCAAGACAGCCACTTCTCT 2288
 DB 411 GGGATTGTGGGCTGGGCTTGGCTTCTAACTGCTGTGTGACCAAGACAGCCACTTCTCT 352
 QY 2289 CCCTAACCTTGTTATGCTTGGCAGCAGTGAGCAGCTGGAGCTAGCGCAACAGTTT 2348
 DB 351 CCCTAACCTTGTTATGCTTGGCAGCAGTGAGCAGCTGGAGCTAGCGCAACAGTTT 292
 QY 2349 GGATTATTGTGTTTTAGATGGAATATTTTTTGTATATAAACTTTATGTGTAACC 2408
 DB 291 GGATTATTGTGTTTTAGATGGAATATTTTTTGTATATAAACTTTATGTGTAACC 232
 QY 2409 CCAATATGAACACTAGATTAAAGGAGTCTCTCTGGTTGAAAGGGAGCTGAGTACCT 2468
 DB 231 CCAATATGAACACTAGATTAAAGGAGTCTCTCTGGTTGAAAGGGAGCTGAGTACCT 172
 QY 2469 CTGGAACTGGAGGCACCTCTGAAAAAGCAAACTGAAACCAAGTGCCTGGGTCACTGTT 2528
 DB 171 CTGGAACTGGAGGCACCTCTGAAAAAGCAAACTGAAACCAAGTGCCTGGGTCACTGTT 112
 QY 2529 ACTCTCTAAGACAGTTTAAAGTGAGACCTGAAAAACATTTTCTTACCTTGAATAGAT 2588
 DB 111 ACTCTCTAAGACAGTTTAAAGTGAGACCTGAAAAACATTTTCTTACCTTGAATAGAT 52
 QY 2589 AGGTTTTTATGTTGGTATATAAGAAATAAAACTAACTTAAACCCCTGAGA 2639

Db 51 AGGTTTTTATGTTGGTATATAAGAAATAAACTAACTATTAAACCTGAAA 1
 RESULT 11
 BM688117
 LOCUS
 DEFINITION
 ACCESSION BM688117
 VERSION
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 489)
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.
 Location/Qualifiers
 FEATURES
 source
 1..489
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-ClO-abx-a-05-0-UI"
 /tissue_type="human retina"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-ClO"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-E-ClO is a cDNA library containing the following
 tissue(s): retina. The library was constructed according
 to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was ligated to an EcoR I adaptor, digested with Not
 I, and cloned directionally into pT73-Pac vector. The
 oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is CCGCG. This library was
 created for the program, Gene Discovery in the Visual
 System, supported by National Eye Institute (NEI)."
 ORIGIN
 Query Match 16.5%; Score 464.8; DB 4; Length 489;
 Best Local Similarity 99.6%; Pred. No. 1.2e-72;
 Matches 466; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2172 TTTAAAGGATGGAGAAAACCTTCTGCTCTCTGGCAAGAAATGGGGAGCTGTGACTGG 2231
 DB 1 TTTAAAGGATGGAGAAAACCTTCTGCTCTCTGGCAAGAAATGGGGAGCTGTGACTGG 60
 QY 2232 ATTGTGGGCTGGGCTTGGCTTCTAACTGCTGTGTGACCCAGACAGCCACTTCTCTCCTCC 2291

```

Db      61  |||||ATTTGGGCTGGCTTGGCTTCTAACTGCTGTGTGACCCCAAGACAGCCACTTCTCTCC 120
QY      2292  |||||TAACCTTGGTTATGCTCTTGGCAGCACAGTGCAGCAGGTGCGACTAGGCGAACAGTCTTTGGA 2351
Db      121  |||||TAACCTTGGTTATGCTCTTGGCAGCACAGTGCAGCAGGTGCGACTAGGCGAACAGTCTTTGGA 180
QY      2352  |||||TTATTGTTTTAGATGCGAATTAATTTTGTGTATATATAAACTCTTATGTGTAACCCCA 2411
Db      181  |||||TTATTGTTTTAGATGCGAATTAATTTTGTGTATATATAAACTCTTATGTGTAACCCCA 240
QY      2412  |||||ATATAGAACTAGATTAAGAGGAGTCTCTCTGTTGGAAGGAGCTGATACCTCTG 2471
Db      241  |||||ATATAGAACTAGATTAAGAGGAGTCTCTCTGTTGGAAGGAGCTGATACCTCTG 300
QY      2472  |||||GAACCTGGAGCAGCTCTGAAAAAGCAAACTGAAAAACAGTGCCTGGGTCACCTGTTACT 2531
Db      301  |||||GAACCTGGAGCAGCTCTGAAAAAGCAAACTGAAAAACAGTGCCTGGGTCACCTGTTACT 360
QY      2532  |||||CCTATAGACAGTTTAAAGTGAGACCTGGAAAAACATTTGCTTTTACCTTGAATAGATAG 2591
Db      361  |||||CCTATAGACAGTTTAAAGTGAGACCTGGAAAAACATTTGCTTTTACCTTGAATAGATAG 420
QY      2592  |||||TTTTTATGTTGGTATATAGAAATTAACCTAACTTAACTTAACTGAGA 2639
Db      421  |||||TTTTTATGTTGGTATATAGAAATTAACCTAACTTAACTTAACTGAGA 468

```

```

RESULT 12
AI424326/c
LOCUS      AI424326      489 bp      mRNA      linear      EST 28-MAR-1999
DEFINITION te90c12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2093974 3',
            mRNA sequence.
ACCESSION  AI424326
VERSION     AI424326.1 GI:4270257
KEYWORDS   Homo sapiens (human)
SOURCE      Homo sapiens
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 489)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
JOURNAL    Contact: Robert Strausberg, Ph.D.
COMMENT    Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 578 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 431.
FEATURES   Location/Qualifiers
            i..489
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:2093974"
            /sex="male"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="NCI CGAP Pr28"
            /notes="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
            with a modified polylinker; Plasmid DNA from the
            normalized library NCI_CGAP_Pr22 was prepared, and as
            circles were made in vitro. Following HAP purification,
            this DNA was used as tracer in a subtractive hybridization

```

```

reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clones IDs
985608-986759, 110192-110195, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo. "
ORIGIN
Query Match      16.4%; Score 461.8; DB 1; Length 489;
Best Local Similarity 99.8%; Pred. No. 4.2e-72;
Matches 463; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
2175  AAGGATGGAGAAACCTTCTGCTCTGCGCAAGAATTGGGGAGCTGTGACTTGGGATT 2234
489   AAGGATGGAGAAACCTTCTGCTCTGCGCAAGAATTGGGGAGCTGTGACTTGGGATT 430
2235  GTGGGCTGGGCTTGGCTTCTAACTGCTGTGTGACCAAGACAGCCACTTCTCTCCCTAA 2294
429   GTGGGCTGGGCTTGGCTTCTAACTGCTGTGTGACCAAGACAGCCACTTCTCTCCCTAA 370
2295  CTTGTTTATGCTTGGCAGCAGTGAGCAGTGCAGCTAGGCGAACAGTCTTCTCTCCCTAA 2354
369   CTTGTTTATGCTTGGCAGCAGTGAGCAGTGCAGCTAGGCGAACAGTCTTCTCTCCCTAA 310
2355  TTGTGTTTTAGATGTTGAATTAATTTTGTGTATATAAACTCTTATGTGTAACCCCAATA 2414
309   TTGTGTTTTAGATGTTGAATTAATTTTGTGTATATAAACTCTTATGTGTAACCCCAATA 250
2415  TAGAACTAGATTAAGAGGAGTCTCTCTGTTGGAAGGAGCTGAGTACCCCTCTGGAA 2474
249   TAGAACTAGATTAAGAGGAGTCTCTCTGTTGGAAGGAGCTGAGTACCCCTCTGGAA 190
2475  CTGGAGCAGCTCTGAAAAAGCAAACTGAAAAACAGTGCCTGGGTCACCTGTTACTCT 2534
189   CTGGAGCAGCTCTGAAAAAGCAAACTGAAAAACAGTGCCTGGGTCACCTGTTACTCT 130
2535  ATAGACAGTCTTAACTGAGACCTGAAAAACATTTGCTTTTACCTTGAATAGATAGTTT 2594
129   ATAGACAGTCTTAACTGAGACCTGAAAAACATTTGCTTTTACCTTGAATAGATAGTTT 70
2595  TTATGTTGTATATAGAAATTAACCTAACTTAACTTAACTGAGA 2639
69   TTATGTTGTATATAGAAATTAACCTAACTTAACTTAACTGAGA 25

```

```

RESULT 13
BM704154
LOCUS      BM704154      475 bp      mRNA      linear      EST 28-FEB-2002
DEFINITION UI-E-CKI-afi-p-12-0-UI-rl UI-E-CKI Homo sapiens cDNA clone
            UI-E-CKI-afi-p-12-0-UI-5', mRNA sequence.
ACCESSION  BM704154
VERSION     BM704154.1 GI:19017412
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 475)
            Bonaldo, M.F., Lennon, G. and Soares, M.B.
            Normalization and subtraction: two approaches to facilitate gene
            discovery
            Genome Res. 6 (9), 791-806 (1996)
JOURNAL    MEDLINE
PUBMED     97044477
COMMENT    8889548
            Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

```

```

Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1. .475
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="UI-E-CK1-afi-p12-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK1"
/note="Organ: eye; Vector: pTT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-CK1 is a normalized cDNA library containing the
following tissue(s): Retina Foveal and Macular. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pTT73-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dr)18 tail. The sequence tag for this
library is GTCC. This library was created for the program,
Gene Discovery in the Visual System, supported by National
Eye Institute (NEI)."
```

ACCESSION	AW020233
VERSION	AW020233.1 GI:5848949
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 455)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-i@mail.nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Seq primer: -40UP from Gibco.
FEATURES	Location/Qualifiers
source	1..455
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:2514063"
	/dev_stage="fetal"
	/lab_host="DH10B (phage-resistant)"
	/clone_lib="Soares thymus NHFTH"
	/notes="Organ: thymus, pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo (dT) primer [5' TGTTCAACTGTGAAGTGCGGGCCGCCAAGCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN	
Query Match	15.9%; Score 448.8; DB 2; Length 455;
Best Local Similarity	99.6%; Pred. No. 8.8e-70;
Matches 450; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
Qy	2188 ACTTCGTCTGCTCTGGCAAGAATTGGGGGACGTGTGACTGGGATTTGGGCTGGGCTT 2247
Db	455 ACTTCGTCTGCTCTGGCAAGAATTGGGGGACGTGTGACTGGGATTTGGGCTGGGCTT 396
Qy	2248 GGCTTCTAACTGCTGTGTGACCAGACGCACCTTCTCCTCCCTAACCTTGGTTATGTC 2307
Db	395 GGCTTCTAACTGCTGTGTGACCCAGACGCACCTTCTCCTCCCTAACCTTGGTTATGTC 336
Qy	2308 TTGGCAGCACAGTGAGCAGTCCGACTAGCGCAACAGTTTTTGGATTATTGTGTTTTAGA 2367
Db	335 TTGGCAGCACAGTGAGCAGTCCGACTAGCGCAACAGTTTTTGGATTATTGTGTTTTAGA 276
Qy	2368 TGTGGAAATTATTTTTTGTATTATAAACTCTTATGTGTAAACCCCAATATAGAACTAGATT 2427
Db	275 TGTGGAAATTATTTTTTGTATTATAAACTCTTATGTGTAAACCCCAATATAGAACTAGATT 216
Qy	2428 AAAAGGAGTCTCTCTGGTTGAAAAGGGAGCTGAGTACCCCTCTGGAACTGGAGGCACCTC 2487
Db	215 AAAAGGAGTCTCTCTGGTTGAAAAGGGAGCTGAGTACCCCTCTGGAACTGGAGGCACCTC 156
Qy	2488 TGA AA AAGCAA ACTGAA AAC CAGTGCCTGGGCTCACTGTTACTCTCTATAAGACAGTTTA 2547
Db	155 TGA AA AAGCAA ACTGAA AAC CAGTGCCTGGGCTCACTGTTACTCTCTATAAGACAGTTTA 96
Qy	2548 AAGTGAACCTGGAAAAACAATTTTGCCTTTACCTTTGAATAGATAGGTTTTTATGTGTGTA 2607
Db	95 AAGTGAACCTGGAAAAACAATTTTGCCTTTACCTTTGAATAGATAGGTTTTTATGTGTGTA 36
Qy	2608 TAAGAAATAAACTAACCTTATTAACCTTGAGA 2639
Db	35 TAAGAAATAAACTAACCTTATTAACCTGAAA 4

RESULT 15
 BM686135
 LOCUS
 DEFINITION 437 bp mRNA linear EST 27-FEB-2002
 UI-E-CKO-aam-b-12-0-UI.r1 UI-E-CKO Homo sapiens cDNA clone
 UI-E-CKO-aam-b-12-0-UI 5', mRNA sequence.
 ACCESSION BM686135
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 437)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers
 1..437
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-CKO-aam-b-12-0-UI"
 /tissue_type="Retina Foveal and Macular"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-CKO"
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-E-CKO is a cDNA library containing the following
 tissue(s): Retina Foveal and Macular. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT7T3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (GT)18 tail. The sequence tag for this library is GTCC.
 This library was created for the program, Gene Discovery
 in the Visual System, supported by National Eye Institute
 (NEI)."

ORIGIN

Query Match 15.5%; Score 436; DB 4; Length 437;
 Best Local Similarity 99.8%; Pred.No. 1.7e-67;
 Matches 436; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1862 GGGCAGCAATAATCTTATAACTTCGGGATGCTGAGGGCGTTTCATGGGAGGACCCCTGG 1921
 Db 1 GGGCAGCAATAATCTTATAACTTCGGGATGCTGAGGGCGTTTCATGGGAGGACCCCTGG 60
 QY 1922 CTTCTCTCTCCCAAGGCATCTTCAACAGTGGTGTCAACAGGAAAAATGGCAGCAATATAC 1981

Db 61 CTTCTCTCTCCCAAGGCATCTTCAACAGTGGTGTCAACAGGAAAAATGGCAGCAATATAC 120
 QY 1982 GTTGACAGGCTGTGGTCTTTTCTGCTTTGAAAGGGTCAGCTGTACTTTAAAGGGGACTGTTTC 2041
 Db 121 GCTGACAGGCTGTGGTCTTTTCTGCTTTGAAAGGGTCAGCTGTACTTTAAAGGGGACTGTTTC 180
 QY 2042 AGCTCTGCTGGTGTGCTGCTGCTGGGACCCCTGCTGCGAACCACCACTCCGCCAACAT 2101
 Db 181 AGCTCTGCTGGTGTGCTGCTGCTGGGACCCCTGCTGCGAACCACCACTCCGCCAACAT 240
 QY 2102 CTTCTCTTTTCCATCCATATATCCCTCAGTATGGACCTTTCCACAACTCCCGACCAATAAGCTGA 2161
 Db 241 CTTCTCTTTTCCATCCATATATCCCTCAGTATGGACCTTTCCACAACTCCCGACCAATAAGCTGA 300
 QY 2162 ATGTTTCTCTTTTAAAGGATGGAGAAAATTTCTGTCTCTCTGGCAAGAAATTTGGGGACT 2221
 Db 301 ATGTTTCTCTTTTAAAGGATGGAGAAAATTTCTGTCTCTCTGGCAAGAAATTTGGGGACT 360
 QY 2222 GTTGACTGGGATTTGGGCTGGGCTTGGCTTCTTAAGTCTGTGTGACCCCAAGACAGCCAC 2281
 Db 361 GTTGACTGGGATTTGGGCTGGGCTTGGCTTCTTAAGTCTGTGTGACCCCAAGACAGCCAC 420
 QY 2282 TTCTCTCTCCCTAACCTT 2298
 Db 421 TTCTCTCTCCCTAACCTT 437

Search completed: June 29, 2005, 05:14:54
 Job time : 6078.87 secs

This Page Blank (uspto)